Multisensory Integration of Visual and Vestibular Signals Improves Heading Discrimination in the Presence of a Moving Object

Kalpana Dokka,¹ Gregory C. DeAngelis,² and Dora E. Angelaki¹

¹Department of Neuroscience, Baylor College of Medicine, Houston, Texas 77030, and ²Department of Brain and Cognitive Sciences, University of Rochester, Rochester, New York 14627

Humans and animals are fairly accurate in judging their direction of self-motion (i.e., heading) from optic flow when moving through a stationary environment. However, an object moving independently in the world alters the optic flow field and may bias heading perception if the visual system cannot dissociate object motion from self-motion. We investigated whether adding vestibular self-motion signals to optic flow enhances the accuracy of heading judgments in the presence of a moving object. Macaque monkeys were trained to report their heading (leftward or rightward relative to straight-forward) when self-motion was specified by vestibular, visual, or combined visual-vestibular signals, while viewing a display in which an object moved independently in the (virtual) world. The moving object induced significant biases in perceived heading when self-motion was signaled by either visual or vestibular cues alone. However, this bias was greatly reduced when visual and vestibular cues together signaled self-motion. In addition, multisensory heading discrimination thresholds measured in the presence of a moving object were largely consistent with the predictions of an optimal cue integration strategy. These findings demonstrate that multisensory cues facilitate the perceptual dissociation of self-motion and object motion, consistent with computational work that suggests that an appropriate decoding of multisensory visual-vestibular neurons can estimate heading while discounting the effects of object motion.

Key words: flow parsing; object motion; optic flow; self-motion; vestibular

Significance Statement

Objects that move independently in the world alter the optic flow field and can induce errors in perceiving the direction of self-motion (heading). We show that adding vestibular (inertial) self-motion signals to optic flow almost completely eliminates the errors in perceived heading induced by an independently moving object. Furthermore, this increased accuracy occurs without a substantial loss in the precision. Our results thus demonstrate that vestibular signals play a critical role in dissociating self-motion from object motion.

Introduction

Accurate judgment of the direction of self-motion (heading) is important for navigation and interaction with objects in the environment. Humans can judge their heading fairly accurately when moving through a stationary environment (Warren et al.,

1988; Telford et al., 1995; Telford and Howard, 1996; Mapstone et al., 2006; Crane, 2012; Cuturi and MacNeilage, 2013). However, observers can exhibit more substantial errors when they judge heading in the presence of objects that move independently in the world (Warren and Saunders, 1995; Royden and Hildreth, 1996, 1999; Mapstone and Duffy, 2010). Although previous psychophysical studies have demonstrated that the visual system is capable of parsing retinal image motion into components related to object motion and self-motion (Warren and Rushton, 2007; Matsumiya and Ando, 2009; Warren and Rushton, 2009a,b; Dokka et al., 2015), the heading errors induced by object motion suggest that the visual system, by itself, cannot fully dissociate object motion from self-motion.

In addition to optic flow, self-motion is generally accompanied by vestibular signals that provide independent information

Received June 12, 2015; revised Aug. 17, 2015; accepted Aug. 21, 2015.

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

This work was supported by National Institutes on Deafness and Other Communication Disorders Grant R03 DC013987 to K.D. and National Institute of Health Grant R01 DC007620 to D.E.A. and Grant R01 EY016178 to G.C.D. We thank Baowang Li for assistance with data collection.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Dora E. Angelaki, Department of Neuroscience, MS: BCM 295, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. E-mail: angelaki@bcm.edu.

DOI:10.1523/JNEUROSCI.2267-15.2015

Copyright © 2015 the authors 0270-6474/15/3513599-09\$15.00/0

about heading. In particular, vestibular signals originating in the otolith organs provide information about translation of the head and are valuable for judging heading (Gu et al., 2007; Angelaki and Cullen, 2008). Although several studies have demonstrated that visual and vestibular signals are integrated to facilitate heading perception (Gu et al., 2008, 2012; Fetsch et al., 2009, 2011; Butler et al., 2010, 2011, 2015; Prsa et al., 2012, 2015; Chen et al., 2013; de Winkel et al., 2013), these studies were performed within an environment that only contained immobile objects. It is currently unknown whether vestibular signals can help mitigate biases in heading perception induced by moving objects. However, recent studies have demonstrated that vestibular signals improve the ability of human observers to judge object motion in the world during self-motion (MacNeilage et al., 2012; Fajen et al., 2013; Dokka et al., 2015). We therefore hypothesized that vestibular signals will enhance the accuracy of heading perception when moving objects are present.

We trained macaques to discriminate heading in the presence or absence of an independently moving object. We find that adding vestibular self-motion signals to optic flow substantially decreases biases in perceived heading that are induced by the moving object. Furthermore, we show that heading discrimination performance still exhibits near-optimal cue integration, even in the presence of object motion. These results demonstrate that vestibular signals can facilitate the dissociation of object motion from self-motion, leading to more accurate heading judgments without substantial loss of sensitivity. Our findings are largely consistent with computational work showing that multisensory visual/vestibular neurons can be decoded to estimate heading while marginalizing over object motion (Kim and DeAngelis, 2010).

Materials and Methods

Experimental subjects. Six male rhesus macaque monkeys (*Macaca mulatta*) weighing ~6–10 kg performed a heading discrimination task. The surgical preparation, experimental apparatus, and methods of data acquisition have been described in detail previously (Gu et al., 2006; Fetsch et al., 2007; Takahashi et al., 2007). Briefly, each animal was chronically implanted with a circular molded, lightweight plastic ring for head restraint, and a scleral coil for monitoring eye movements inside a magnetic field (CNC Engineering). Behavioral training was accomplished using standard operant conditioning. All surgical and experimental procedures were approved by the Institutional Animal Care and Use Committees at Washington University in St. Louis and Baylor College of Medicine in accordance with National Institutes of Health guidelines.

Apparatus. A custom-built experimental apparatus was used to test the monkeys on a heading discrimination task (Gu et al., 2007, 2008; Fetsch et al., 2009, 2011). The apparatus consisted of two main components: (1) a 6 degree-of-freedom motion platform (6DOF 2000E, Moog) that provided inertial motion stimuli; and (2) a virtual reality system, consisting of a three-chip DLP projector (Christie Digital Mirage 2000) and a rear projection screen, that provided visual stimuli. The projector and screen were rigidly mounted on the motion platform and did not move relative to the platform. Monkeys sat in a primate chair mounted on top of the motion platform and inside the magnetic field coil frame. Each monkey's head was firmly restrained and was positioned such that the horizontal stereotaxic plane of the head was earth-horizontal, with the axis of rotation passing through the center of the head (i.e., the midline point along the interaural axis). The distance from the monkey's eyes to the display screen was 32 cm. The sides and top of the apparatus were covered with a black matte material such that the only visual motion experienced by a fixating animal was that projected onto the display screen. Passive inertial translation of the monkey in the horizontal plane followed a smooth Gaussian velocity profile with a peak velocity of 0.35 m/s (displacement = 13 cm, 1 s duration, peak acceleration = 1.4 m/s²).

Visual stimuli. Visual stimuli depicted movement of the animal through a 3D cloud of stars that occupied a virtual space 100 cm wide, 100 cm tall, and 40 cm deep. Given that physical and technical constraints of our apparatus limit the speeds of self-motion that we can present and the viewing distance, we use a rather near and shallow 3D cloud of stars. If we had simulated a more distant and deep 3D cloud, then the retinal speeds would be low and binocular disparities and star densities (on the display) would become rather large. With the 3D volume we use, the range of retinal motion speeds of the stars is robust (maximum retinal speed of 54°/s for the nearest stars at the most eccentric rightward/leftward headings) and the range of binocular disparities is more balanced around the plane of fixation.

At the beginning of each trial, the virtual star cloud was centered in depth on the plane of the display screen. Thus, the nearest edge of the star cloud was 12 cm in front of the animal at the start of each trial. Each star in the 3D cloud was a triangle having a base width of 0.15 cm and a height of 0.15 cm. Star density was 0.01 stars/cm³, resulting in ~4000 stars filling the 3D cloud. The star cloud was rendered stereoscopically as a red/green anaglyph, and the animal viewed the stimuli through Kodak Wratten filters (red #29 and green #61). With these filters, binocular cross talk was 1.7% for the green channel and 2.3% for the red channel. As these filters were passive, they did not require any synchronization with the projector.

Naturalistic simulations of self-motion were generated using the OpenGL graphics libraries, along with an accelerated graphics board (NVIDIA Quadro 2000). The 3D star cloud represented an immobile virtual world through which self-motion was simulated by moving the OpenGL camera along the exact trajectory of translation of the animal. As the animal moved through this virtual environment, stars in the 3D cloud moved with a velocity pattern that precisely simulated the animal's direction and speed of translation. The forward headings probed in this study were simulated by an expanding radial flow pattern with a focus of expansion that varied systematically with heading.

Because the 3D star cloud was generated by drawing triangles in a virtual environment and moving the OpenGL camera through it, the visual display also contained a variety of 3D cues, including horizontal disparity, motion parallax, star size, and star density. These cues accurately simulated motion of the observer relative to the 3D star field. For example, during simulated forward motion, the size of each star in the image increases over time, and the binocular disparity of each star changes to reflect the distance of the animal to that star. Similarly, stars located closer to the monkey move faster in the image than distant stars, providing motion parallax information. Correspondingly, star density within the image changed as the animal moved through the 3D volume, with density decreasing for forward headings as dictated by movement of the OpenGL camera through the world-fixed cloud.

Monkeys experienced three self-motion conditions: (1) a "vestibular" condition in which inertial motion of the platform was presented without optic flow (only fixation and choice targets were displayed on the projection screen); (2) a "visual" condition in which the platform remained stationary while optic flow simulated motion of the observer through a 3D cloud of stars; and (3) a "combined" condition in which both inertial motion and optic flow were presented congruently and simultaneously. The visual stimulus sequence for the visual and combined conditions was identical and simulated translation of the animal along a smooth Gaussian velocity profile (displacement = 13 cm, duration = 1 s).

In most trials, the visual display also contained a single spherical object (radius 10°) that moved rightward or leftward in the virtual world. The object was composed of the same red and green stars as the background 3D cloud. Star density of the object was 1 star/cm³ resulting in ~520 stars forming the object. The object was located in depth at the center of the 3D star cloud such that the object was 32 cm in front of the animal at the start of each trial. When the animal moved forward, the object became closer to the animal over time such that it also became larger in the image. As the star density of the object was much greater than that of the 3D star cloud, the object was easily segmented from the background optic flow. Moreover, the object was transparent such that it did not occlude portions of the 3D

Figure 1. Illustration of observer–object interaction. The monkey moves forward in the world (v_{self}) while the object independently moves rightward in the world (v_{obj}). The object motion as seen by the monkey's eye ($v_{\text{obj-eye}}$) represents a vector sum of the object motion in the world (v_{obj}) and the motion of the world relative to the eye that is caused by self-motion ($v_{\text{w-self}}$; this vector is equal and opposite to v_{self}). Object motion relative to the (moving) eye ($v_{\text{obj-eye}}$), thus implicitly contains information about the monkey's self-motion.

Figure 2. Example data. Psychometric functions of a representative monkey from a single experimental session performed at the matching visual motion coherence. Data are shown for no object motion (**A**), for rightward object motion (**B**), and for leftward object motion (**C**). Colors represent the different self-motion conditions: vestibular (blue), visual (green), and combined (red). The bias in perceived heading (point of subjective equality) is given by the heading at which a curve crosses 0.5 on the vertical axis.

star cloud. Background stars could appear in front of, behind, or within the object.

Object motion followed a smooth Gaussian velocity profile with a peak velocity of 0.3 m/s (displacement = 10 cm, 1 s duration, peak acceleration = 1.07 m/s²). Similar to the experimental paradigm used by Royden and Hildreth (1996), the motion of the object in the world was chosen to lie within the frontoparallel plane (Fig. 1). For each trial, the initial location of the object depended on the subsequent direction of object motion. For rightward object motion, the object was initially located 17° to the left of the head-centered fixation target. The object moved rightward such that it came to a stop at the location of the fixation point at the end of the trial. Similarly, for leftward object motion, the object was initially located 17° to the right of the fixation point and moved leftward such that it came to a stop at the fixation target location. The movement of the object was temporally synchronized with the monkey's self-motion in all trials. In total, monkeys experienced three object motion conditions: no-object, rightward motion (in the world), and leftward motion; these three object motion conditions were crossed with the three self-motion conditions (vestibular, visual, and combined).

Although the object only moved laterally in the world within a frontoparallel plane (Fig. 1, left panel; v_{obj}), the motion of the object as viewed by the monkey's eye (Fig. 1, right panel; $v_{\text{obj-eye}}$) is a vector sum of object motion in the world (v_{obj}) and the motion of the world relative to the eye (Fig. 1, right panel; $v_{\text{w-self}}$) that results from self-motion (v_{self}). Specifically, the horizontal component of object motion in the image is a combination of object motion in the world and the horizontal component of

the monkey's self-motion. Moreover, the looming motion of the object during forward headings is related to the forward component of self-motion. As a result, object motion relative to the eye ($v_{\text{obj-eye}}$) implicitly contains information about self-motion. Thus, during vestibular self-motion in the presence of a moving object, in addition to vestibular signals, the retinal motion of the object provided a visual cue to heading even though there was no background optic flow. Therefore, unlike the vestibular condition without a moving object, the vestibular condition with object motion was not a purely unimodal heading stimulus but indeed provided both visual and vestibular cues to self-motion.

Heading discrimination task. Monkeys performed a single-interval, two-alternative forced-choice heading discrimination task. In each trial, the monkey was first required to fixate a central target (0.2° diameter) for 200 ms before stimulus onset (fixation windows spanned 2° × 2° visual angle). The fixation target remained head-fixed throughout the self-motion trajectory (as if attached to the windshield of a car, for example) such that no eye movements were required to maintain visual fixation during real or simulated self-motion. If fixation was broken at any time during the stimulus presentation, the trial was aborted and the data were discarded. During the trial, the monkey was presented with a real and/or simulated translational motion in the horizontal plane, with or without object motion. The monkey was required to indicate his perceived heading relative to straight-forward by making a saccade to one of two choice targets illuminated at the end of the trial. Choice targets were presented only after the last frame of the heading stimulus ended. Monkeys were rewarded with a drop of juice when they reported heading correctly.

Heading was varied in fine steps around straight-forward (0°). Three monkeys were first tested with logarithmically spaced heading angles: ±16°, ±8°, ±4°, ±2°, and ±1°, similar to previous studies of heading discrimination in the absence of moving objects (Gu et al., 2008). Data from these monkeys revealed significant biases in perceived heading when object motion was present (e.g., Fig. 2B,C). However, with large biases (large shifts of the point of subjective equality), logarithmically spaced heading angles may yield unreliable discrimination thresholds (or Just-Noticeable-Differences) due to the nonuniform spacing of the tested headings. Therefore, two additional monkeys that were naive to this experiment, as well as one animal that was previously tested with logarithmically spaced headings, performed the heading discrimination task with quasi-linearly spaced headings: ±20°, ±16°, ±12°, ±8°, ±4°, and ±2°. Positive values indicate headings to the right of straight-forward, whereas negative values indicate headings to the left of straight-forward. Each heading was presented 10 times for each of the 9 stimulus conditions. Thus, in a single block of trials, monkeys performed a total of either 900 trials (10 logarithmically spaced headings × 3 self-motion conditions × 3 object motion conditions × 10 repetitions) or 1080 trials (12 quasi-linearly spaced headings × 3 self-motion conditions × 3 object motion conditions × 10 repetitions). All self-motion conditions, object motion conditions, and headings were randomly interleaved within a single block of trials for one session.

Reliability of visual self-motion cues was varied by manipulating the motion coherence of optic flow (Fetsch et al., 2009). Briefly, a fraction of stars corresponding to the percentage coherence was chosen randomly

on each display update. These “signal” stars changed location according to the optic flow field associated with the heading; the remaining “noise” stars were randomly relocated within the 3D star cloud. Depth cues associated with a star were always accurate based on its current location within the 3D cloud, for both signal and noise stars. Each monkey performed the heading discrimination task at three levels of visual motion coherence: “high,” “matching,” and “low.” At high coherence, which was fixed at 80% for all monkeys, visual heading discrimination was more precise than vestibular heading discrimination (i.e., visual thresholds were smaller than vestibular thresholds). The matching coherence for each monkey was chosen such that the visual heading thresholds were approximately equal to vestibular thresholds. Across monkeys, the matching coherence varied from 11% to 47%. At the low coherence, visual heading thresholds were substantially greater than vestibular thresholds, and low coherence values varied from 5% to 19%. Coherence was fixed in each block of trials and was pseudo-randomly varied across blocks; each coherence was tested in three blocks of trials. The motion coherence of the star pattern forming the object was always fixed at 100%.

Data analysis. For each animal, motion coherence, and combination of self-motion and object motion, a psychometric function was constructed by calculating the proportion of rightward choices as a function of heading. These data were fit with a cumulative Gaussian function using the *psignifit* toolbox, a MATLAB software package (The MathWorks) that implements the maximum-likelihood method of Wichmann and Hill (2001a,b). These fits yielded two parameters that characterize the accuracy and precision of heading perception: bias and threshold. Bias (or point of subjective equality) corresponds to the heading which yields 50% rightward and 50% leftward choices. A bias close to 0° indicates accurate heading discrimination. Threshold is given by the SD (σ) of the cumulative Gaussian fit and represents the inverse precision of heading discrimination. A small threshold indicates highly precise (i.e., sensitive) heading discrimination.

In addition to measuring the psychophysical threshold during combined visual-vestibular self-motion, we predicted the threshold expected if visual and vestibular signals are integrated optimally. Optimal cue integration theory predicts that heading performance with combined vestibular-visual signals will be more precise (i.e., smaller threshold) than when self-motion is specified by visual or vestibular cues (Gu et al., 2008, 2012; Fetsch et al., 2009, 2011; Butler et al., 2010, 2011, 2015; Prsa et al., 2012, 2015; Chen et al., 2013; de Winkel et al., 2013). Specifically, the predicted threshold in the combined condition, σ_{com} , can be calculated as follows:

$$\sigma_{\text{com}} = \sqrt{\frac{\sigma_{\text{vest}}^2 \sigma_{\text{vis}}^2}{\sigma_{\text{vest}}^2 + \sigma_{\text{vis}}^2}}$$

where σ_{vest} and σ_{vis} represent the single-cue visual and vestibular heading thresholds. Thus, the single-cue thresholds measured in the absence of object motion were used to predict the combined threshold without object motion. To summarize the optimality of cue integration, we computed a ratio of the empirically measured threshold in the combined condition to the predicted threshold. A threshold ratio close to 1 indicates that visual and vestibular signals are combined near-optimally.

We also sought to characterize the optimality of visual-vestibular cue integration in the presence of the moving object. However, the retinal motion of the object reflects the vector sum of object motion in the world and the monkey's self-motion (Fig. 1). Thus, the vestibular condition with object motion includes some visual information about self-motion. As a result, it would be incorrect to apply optimal cue integration theory to the visual and vestibular conditions with object motion. To avoid this problem, we predicted the combined threshold from the combination of the vestibular threshold without object motion and the visual threshold with object motion.

Data analyses were performed using MATLAB and Minitab. For the analysis of bias in perceived heading, data from all six monkeys were used. For the analysis of thresholds, data from the three monkeys tested with quasi-linearly spaced headings were used. For statistical analysis, data from rightward and leftward object motion were grouped together.

Specifically, the bias measured with leftward object motion was multiplied by -1 and grouped along with the bias for rightward object motion (Dokka et al., 2015). A repeated-measures ANOVA with three factors (object motion, self-motion, visual motion coherence) was performed on both the bias and threshold values. When the ANOVA revealed a significant effect of a factor, Tukey–Cramer *post hoc* multiple comparisons were performed to identify groups with significant differences. To test whether visual and vestibular signals were integrated optimally, a *t* test compared the threshold ratio with 1.

Results

We trained rhesus monkeys to perform a heading discrimination task based on vestibular, visual, and combined (vestibular-visual) self-motion cues. Animals performed this task in the absence of moving objects and in the presence of an object that moved rightward or leftward in the world. We sought to determine how a moving object would bias heading percepts and whether the combination of visual and vestibular self-motion cues would ameliorate biases induced by the moving object. We also varied the reliability (i.e., motion coherence) of optic flow to examine how the interaction between self-motion and object motion depends on the quality of the self-motion cues.

Example experiment

Example psychometric functions from one session of testing in a representative animal are shown in Figure 2 (data are presented for the “matching” motion coherence; see Materials and Methods). In the absence of object motion (Fig. 2A), the monkey accurately perceived his heading, as indicated by the point of subjective equality (PSE, or bias) lying close to 0° for all self-motion conditions. In addition, the precision of heading discrimination was greater (steeper slope) in the combined condition, as shown previously (Fetsch et al., 2009). In contrast, object motion induced clear biases in heading perception. Rightward object motion induced substantial leftward heading biases in both the vestibular (Fig. 2B, blue curve) and visual (Fig. 2B, green curve) conditions, as quantified by PSE values of 4.52° for vestibular self-motion and 5.88° for visual self-motion. Analogously, leftward object motion induced rightward heading biases in the vestibular (-5.53°) and visual (-5.16°) conditions (Fig. 2C). Thus, consistent with previous findings for visual self-motion (Warren and Saunders, 1995), a moving object induced systematic biases in heading perception when self-motion was indicated by only visual or vestibular cues. Critically, however, the combination of visual and vestibular self-motion cues greatly reduced the bias induced by object motion (Fig. 2B,C, red curves), with PSE values of 0.63° and -2.05° for rightward and leftward object motion, respectively.

Summary of bias effects

Similar effects of object motion on heading perception were observed across sessions and animals (Fig. 3). Heading perception in the no-object condition was generally accurate, as indicated by biases close to 0° for all self-motion conditions and motion coherences (Fig. 3, black curves). In contrast, rightward and leftward object motion consistently induced biases of opposite sign in heading perception (Fig. 3, orange curves). The main effect of object motion condition on heading bias was highly significant (repeated-measures ANOVA: $F_{(1,580)} = 152.85$, $p < 0.0001$; see Materials and Methods). Furthermore, a significant interaction effect ($F_{(2,580)} = 27.2$, $p < 0.001$) between self-motion and object motion conditions revealed that object motion biases differed across self-motion conditions. Specifically, with a moving object,

Figure 3. Summary of biases in perceived heading. Heading bias, averaged across sessions and monkeys, is shown for stimulus conditions without object motion (black symbols), with rightward object motion (filled orange symbols), and with leftward object motion (open orange symbols). Points along the abscissa in each panel indicate the different self-motion conditions: vestibular (Vest), visual (Vis), and combined (Com). Different panels represent results obtained for high (A), matching (B), and low (C) visual motion coherences. Dashed horizontal line indicates zero bias. Error bars indicate SEM.

Figure 4. Summary of heading biases for individual animals. Comparison of biases measured during the combined condition with biases measured during the vestibular (A) and visual (B) conditions, in the presence of a moving object. Each symbol shape represents data from an individual monkey, averaged across three blocks of trials. Filled symbols represent data collected with rightward object motion. Open symbols represent data collected with leftward object motion. Blue, green, and red colors represent high, matching, and low motion coherences. Dashed horizontal line indicates zero bias. Dashed diagonal line indicates unity slope. Error bars indicate SEM.

the bias in the combined condition was significantly less than biases in the vestibular (Tukey–Cramer *post hoc* comparison; $p < 0.001$) and visual ($p < 0.001$) conditions (data pooled across coherences), but not significantly different from the bias in the no-object combined condition (Tukey–Cramer *post hoc* comparison: $p = 0.33$). Thus, biases induced by object motion were almost completely eliminated by multisensory cues in the combined condition.

Data from individual animals were largely consistent with the pattern of results described above. Figure 4, A and B, compares biases in the combined condition with biases in the vestibular and visual conditions, respectively, in the presence of a moving object. The majority of the filled symbols (corresponding to rightward object motion) lie in the upper right quadrant, whereas the majority of the open symbols (leftward object motion) lie in the lower left quadrant. Leftward object motion data from one animal are outliers in Figure 4A. Importantly, the slope of the relationship between combined and vestibular biases was significantly less than unity (Fig. 4A; slope = 0.20, 95% CI = 0.19, 0.21), as was the slope of the relationship between combined and visual biases (Fig. 4B; slope = 0.23, 95% CI = 0.20, 0.25). These results show clearly that multisensory combination of vestibular and visual self-motion cues improves the accuracy of heading perception in the presence of a moving object.

We also explored how the interaction between object motion and self-motion was modulated by visual motion coherence. Although there was no significant main effect of coherence on heading biases (repeated-measures ANOVA; $F_{(2,580)} = 0.93$, $p = 0.39$), there was a significant interaction between coherence, object motion, and self-motion conditions ($F_{(4,580)} = 4.44$, $p = 0.002$). In the no-object condition, biases did not depend significantly on coherence in any of the self-motion conditions (Tukey–Cramer *post hoc* comparisons; $p > 0.8$). However, in the presence of object motion, coherence significantly modulated bias in the visual condition such that it was significantly smaller at high coherence than matching coherence (Tukey–Cramer *post hoc* comparisons; $p = 0.008$) and significantly greater at low coherence than matching coherence ($p = 0.03$). Thus, when self-motion was specified by visual cues alone, the influence of object motion on heading accuracy varied inversely with the reliability of the visual self-motion. This suggests that the visual system has greater trouble in dissociating object motion from self-motion when background visual motion cues are unreliable.

Summary of threshold effects

Although our main focus has been on how object motion affects the accuracy of heading percepts, it is also of considerable interest to know how object motion affects the precision of heading discrimination and the optimality of cue integration.

Overall, we find a significant main effect

of object motion on heading thresholds (Fig. 5; repeated-measures ANOVA; $F_{(1,277)} = 33.9$, $p < 0.001$) such that the presence of object motion reduces heading thresholds and improves precision. We also find a significant interaction between object motion and self-motion condition ($F_{(2,277)} = 8.08$, $p < 0.001$), which reflects the observation that object motion significantly reduces heading discrimination thresholds for the vestibular and visual conditions (Tukey–Cramer *post hoc* comparisons; $p = 0.001$ for vestibular; $p < 0.001$ for visual condition) but does not significantly affect thresholds in the combined condition ($p = 0.5$). We suggest that object motion increases the precision of heading perception because the retinal motion of the object is a vector sum of object motion and observer motion. Thus, in trials with object motion, monkeys have an additional visual cue to self-motion provided by one component of retinal object motion, in addition to vestibular or background optic flow signals. Indeed, it is worth noting that thresholds for the vestibular condition with object motion are fairly similar to those in the combined condition without object motion (Fig. 5). Thus, animals can use the self-motion component of object motion to improve heading sensitivity, but because they cannot completely dissociate the retinal motion components, object motion in the world biases heading percepts in the absence of vestibular self-motion signals.

Consistent with previous work (Fetsch et al., 2009), we found that heading thresholds showed a significant interaction between coherence and self-motion condition (repeated-measures ANOVA: $F_{(4,277)} = 51.00$, $p < 0.001$) such that visual heading thresholds declined with increasing motion coherence, whereas vestibular and combined thresholds did not depend significantly on coherence. However, we found no significant interaction between coherence and object motion condition ($F_{(2,277)} = 0.61$, $p = 0.54$), indicating that the effect of coherence on heading thresholds was similar whether or not there was a moving object present.

Our main interest in examining heading thresholds was to determine whether object motion alters the optimality of visual-vestibular cue integration. In the absence of a moving object, our subjects exhibited greater precision of heading discrimination in the combined condition. The combined threshold was significantly less than thresholds for the vestibular (Tukey–Cramer *post hoc*: $p < 0.001$) and visual ($p < 0.001$) conditions (Fig. 5) when no moving object was present. Moreover, the ratio of the combined threshold to the optimal prediction (see Materials and Methods) was not significantly different from unity (Fig. 6A, black bar; t test, $p = 0.51$). A threshold ratio close to unity indicates that the observer implemented a cue combination strategy that took into account the reliability of the unimodal visual and vestibular cues. Thus, we replicate previous findings (Fetsch et al., 2009) regarding visual-vestibular cue integration in the absence of object motion.

To examine optimality of cue integration in the presence of a moving object, it is not appropriate to predict the combined threshold from vestibular and visual thresholds measured in the presence of a moving object, as object motion would then contribute twice to the predicted threshold (see Materials and Methods). Rather, we predicted the combined threshold from the visual threshold measured with object present and the vestibular threshold measured in the absence of the moving object. Consistent with some degree of cue integration, combined thresholds are generally less than visual thresholds in the presence of object motion (Fig. 6B; $p < 0.001$). Moreover, combined thresholds with object motion are significantly less than vestibular thresholds measured without object motion (Tukey–Cramer *post hoc* comparison: $p < 0.001$, data not shown).

We then computed the ratio of combined to predicted thresholds in the presence of object motion, and we found that this ratio is not significantly greater than unity (Fig. 6A, orange bars; t test, $p = 0.11$). Moreover, a paired t test on the threshold ratios with and without object motion did not show a significant difference ($t_{(63)} = -1.34$, $p = 0.18$, data pooled across leftward and rightward object motion). Although threshold ratios tend to be slightly greater than unity with object motion, our data do not allow us to reject the null hypothesis that cue integration remains optimal in the presence of a moving object. Thus, the neural computations by which vestibular signals are used to attenuate biases caused by object motion appear to largely preserve the benefits of cue integration for heading discrimination.

Discussion

When moving observers view objects that move independently in the world, retinal image motion reflects a vector sum of object

Figure 5. Summary of heading discrimination thresholds. Mean heading discrimination thresholds, averaged across three monkeys tested with the quasi-linear range of headings (see Materials and Methods) during vestibular (Vest), visual (Vis), and combined (Com) self-motion. Different panels represent results obtained for high (A), matching (B), and low (C) visual motion coherences. Error bars indicate SEM.

motion and self-motion. To accurately judge heading, observers must be able to dissociate object motion from self-motion. However, the visual system by itself cannot completely separate these two components as demonstrated by the substantial heading errors observed when self-motion is specified by visual cues alone. Our results demonstrate that adding congruent vestibular signals brings about a much more complete dissociation of object motion from self-motion, leading to largely unbiased heading perception. These findings, together with complementary recent work on object motion perception (Fajen et al., 2013; Dokka et al., 2015), suggest a critical role for vestibular signals (possibly together with somatosensory cues that are also present during inertial motion) in the important process of parsing visual image motion into components related to self-motion and object motion. Our results are consistent with a recent proposal (Kim and DeAngelis, 2010) that multisensory neural representations of self-motion allow efficient marginalization computations that discount the effects of object motion, as discussed below.

Biases in visual heading perception

Consistent with our results for the visual condition, previous studies (Warren and Saunders, 1995; Royden and Hildreth, 1996, 1999) also reported errors in visual heading perception in the presence of a moving object. Specifically, our findings are in agreement with those of Warren and Saunders (1995) who observed heading biases in the direction opposite to the object motion. However, Royden and Hildreth (1996) reported heading biases in the direction of object motion. This difference may be due to the specific characteristics of the object motion stimulus. In the present study as well as the paradigm tested by Warren and Saunders (1995), the observer approached the object such that the distance between object and observer changed during the trial. In contrast, the object always moved in the frontoparallel plane at a fixed distance from the observer in the paradigm tested by Royden and Hildreth (1996). Thus, approaching objects may bias heading differently than objects that always move a fixed distance. In addition, these previous studies found that heading perception was significantly biased when a moving object obscured the focus of expansion (Royden and Hildreth, 1996) or when the object crossed the observer's path of motion (Warren and Saunders, 1995), whereas no substantial biases were observed otherwise.

Numerous methodological differences between our experiments and those of previous studies hamper a direct quantitative comparison of results. Our goal was to investigate the influence of multisensory cues on heading perception in the presence of a moving object, rather than to explore how various nuances of the

Figure 6. Multisensory improvement in heading discrimination thresholds. **A**, Threshold ratio, defined as the ratio of measured to predicted thresholds during combined visual-vestibular self-motion, is shown for conditions without object motion (black bar), with rightward object motion (filled orange bar), and leftward object motion (hatched orange bar). Data from all coherences have been combined together. Dashed horizontal line indicates a threshold ratio of unity. Error bars indicate SEM. **B**, Comparison of thresholds measured in the combined and visual conditions, in the presence of a moving object. Each symbol shape represents data from an individual monkey, averaged across three blocks of trials. Filled symbols represent rightward object motion. Open symbols represent leftward object motion. Blue, green, and red represent data from high, matching, and low motion coherences, respectively. Dashed diagonal line indicates unity slope.

visual display influence heading perception. Among the notable differences, the visual stimuli in our experimental setup presented both rightward and leftward headings, unlike Royden and Hildreth (1996) who only presented rightward headings; we presented headings that were defined in world-coordinates, unlike Warren and Saunders (1995) who defined heading with respect to a “probe” stimulus; we trained monkeys to make heading choices relative to an internal reference of straight-forward, unlike Warren and Saunders (1995) whose human subjects indicated heading choices with respect to a “probe” stimulus having a variable location; and we fixed the initial location of the object for a given direction of object motion, unlike Warren and Saunders (1995) and Royden and Hildreth (1996). Despite these various differences in experimental design, our results are broadly consistent with those of previous studies and reinforce the conclusion that the visual system, on its own, cannot completely dissociate object motion from self-motion.

Biases in vestibular heading perception

In the present experiment, monkeys exhibited significant heading biases when they viewed object motion during vestibular (inertial) self-motion. To our knowledge, previous studies have not reported such biases in vestibular heading perception induced by a moving object. These biases suggest that monkeys did not fully dissociate object motion from self-motion, even in the presence of vestibular self-motion signals.

This may be surprising as one might expect that the presence of vestibular signals, which provide an independent cue to self-motion, should bring about a complete dissociation of object motion from self-motion. For example, the monkey could accurately judge heading in the vestibular condition just by ignoring the visual stimuli altogether. However, the heading biases induced by object motion in the vestibular condition may partly reflect the prior behavioral training these monkeys received. Each of these animals was previously trained to perform heading discrimination (in the absence of object motion) with vestibular, visual, and combined self-motion cues. As a result of this training, the monkeys may have expected object motion to carry

information about self-motion and therefore incorporated it in their heading estimates. This view is strengthened by the similarities in heading biases measured in the presence of a moving object during vestibular and visual self-motion.

Potential neural mechanisms for dissociating object motion and self-motion

Previous studies have proposed visual mechanisms, such as template matching (Warren and Saunders, 1995), differential motion operators (Hildreth, 1992; Royden, 2002), motion pooling in MT⁺ along with template matching (Layton et al., 2012), and visual segmentation (Raudies and Neumann, 2013), to explain the influence of object motion on visual heading perception. Whereas these proposed visual mechanisms can explain the influence of a moving object when self-motion is specified by optic flow alone, our finding that vestibular signals reduce heading biases indicates the existence of comple-

mentary multisensory mechanisms that contribute to dissociating self-motion and object motion.

Our findings have important implications for the underlying neural mechanisms that allow us to estimate heading in the presence of moving objects. We suggest that multisensory improvements in the accuracy of heading perception in the presence of object motion are due to a selective decoding of specialized multisensory neurons that are sensitive to vestibular self-motion, optic flow, and object motion. For example, neurons in the dorsal subdivision of the medial superior temporal area (MSTd) have been shown to exhibit selectivity to inertial motion in darkness (Duffy, 1998; Bremmer et al., 1999; Page and Duffy, 2003; Gu et al., 2006; Gu et al., 2007; Takahashi et al., 2007), optic flow (Tanaka et al., 1986, 1989; Tanaka and Saito, 1989; Duffy and Wurtz, 1991; Duffy, 1998; Gu et al., 2008; Fetsch et al., 2011), and object motion (Logan and Duffy, 2006). Importantly, these multisensory MSTd neurons come in two types: congruent cells and opposite cells (Gu et al., 2006, 2008). Congruent cells have similar vestibular and visual heading preferences and appear to be well suited to mediating multisensory cue integration and reliability-based cue weighting for heading perception (Gu et al., 2008; Fetsch et al., 2011). In contrast, opposite cells tend to have highly dissimilar (often 180° apart) vestibular and visual heading preferences and are not beneficial for cue integration or cue weighting. Similar findings for congruent and opposite cells have also been observed in the ventral intraparietal area (Chen et al., 2013). Thus, the functional role of opposite cells has remained unclear, and we suggest that opposite cells play an important role in dissociating self-motion and object motion.

Specifically, we have recently shown that selectively decoding a mixed population of congruent and opposite neurons can lead to heading estimates that are substantially more robust to object motion (Kim and DeAngelis, 2010). In this scheme, both congruent and opposite cells are decoded according to their vestibular heading preferences, thereby largely cancelling the biases that are induced by object motion in the absence of vestibular input. We have shown that this decoding strategy is predicted by learning a linear transforma-

tion of responses that is optimized to estimate heading by marginalizing out the effect of the object motion (Kim and DeAngelis, 2010). Marginalization, a key operation by which the effects of “nuisance” parameters (e.g., object motion in our case) can be mitigated, has been previously described in motor control (Wolpert et al., 1995) and visual object tracking (Körding, 2007). Our recent computational work (Kim and DeAngelis, 2010) suggests that congruent and opposite cells play key roles in marginalizing over object motion to compute heading. This mechanism may account for the behavioral effects of vestibular signals on heading biases that were observed in the present study. Moreover, preliminary results from recordings in area MSTd suggest that applying this marginalization approach to real neural data can predict reductions in heading biases due to multisensory integration (Sasaki et al., 2012).

In conclusion, our behavioral results indicate clearly that incorporating vestibular signals into computations of self-motion renders heading estimates substantially more robust to the effects of object motion. These findings are consistent with a recent proposal (Kim and DeAngelis, 2010) that a diverse population of multisensory neurons provides a sufficient substrate to perform marginalization operations that largely eliminate the effect of object motion on heading estimates. Thus, our findings may reflect a general mechanism by which multisensory integration allows the brain to tease apart the sources of sensory inputs.

References

- Angelaki DE, Cullen KE (2008) Vestibular system: the many facets of a multimodal sense. *Annu Rev Neurosci* 31:125–150. [CrossRef Medline](#)
- Bremmer F, Kubischik M, Pökel M, Lappe M, Hoffmann KP (1999) Linear vestibular self-motion signals in monkey medial superior temporal area. *Ann N Y Acad Sci* 871:272–281. [CrossRef Medline](#)
- Butler JS, Smith ST, Campos JL, Bühlhoff HH (2010) Bayesian integration of visual and vestibular signals for heading. *J Vis* 10:23. [CrossRef Medline](#)
- Butler JS, Campos JL, Bühlhoff HH, Smith ST (2011) The role of stereo vision in visual-vestibular integration. *Seeing Perceiving* 24:453–470. [CrossRef Medline](#)
- Butler JS, Campos JL, Bühlhoff HH (2015) Optimal visual-vestibular integration under conditions of conflicting intersensory motion profiles. *Exp Brain Res* 233:587–597. [CrossRef Medline](#)
- Chen A, DeAngelis GC, Angelaki DE (2013) Functional specializations of the ventral intraparietal area for multisensory heading discrimination. *J Neurosci* 33:3567–3581. [CrossRef Medline](#)
- Crane BT (2012) Direction specific biases in human visual and vestibular heading perception. *PLoS One* 7:e51383. [CrossRef Medline](#)
- Cuturi LF, MacNeilage PR (2013) Systematic biases in human heading estimation. *PLoS One* 8:e56862. [CrossRef Medline](#)
- de Winkel KN, Soya F, Barnett-Cowan M, Bühlhoff HH, Groen EL, Werkhoven PJ (2013) Integration of visual and inertial cues in the perception of angular self-motion. *Exp Brain Res* 231:209–218. [CrossRef Medline](#)
- Dokka K, MacNeilage PR, DeAngelis GC, Angelaki DE (2015) Multisensory self-motion compensation during object trajectory judgments. *Cereb Cortex* 25:619–630. [CrossRef Medline](#)
- Duffy CJ (1998) MST neurons respond to optic flow and translational movement. *J Neurophysiol* 80:1816–1827. [Medline](#)
- Duffy CJ, Wurtz RH (1991) Sensitivity of MST neurons to optic flow stimuli: II. Mechanisms of response selectivity revealed by small-field stimuli. *J Neurophysiol* 65:1346–1359. [Medline](#)
- Fajen BR, Parade MS, Matthis JS (2013) Humans perceive object motion in world coordinates during obstacle avoidance. *J Vis* 13:pii25. [CrossRef Medline](#)
- Fetsch CR, Wang S, Gu Y, DeAngelis GC, Angelaki DE (2007) Spatial reference frames of visual, vestibular, and multimodal heading signals in the dorsal subdivision of the medial superior temporal area. *J Neurosci* 27:700–712. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE (2011) Neural correlates of reliability-based cue weighting during multisensory integration. *Nat Neurosci* 15:146–154. [CrossRef Medline](#)
- Gu Y, Watkins PV, Angelaki DE, DeAngelis GC (2006) Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J Neurosci* 26:73–85. [CrossRef Medline](#)
- Gu Y, DeAngelis GC, Angelaki DE (2007) A functional link between area MSTd and heading perception based on vestibular signals. *Nat Neurosci* 10:1038–1047. [CrossRef Medline](#)
- Gu Y, Angelaki DE, DeAngelis GC (2008) Neural correlates of multisensory cue integration in macaque MSTd. *Nat Neurosci* 11:1201–1210. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Hildreth EC (1992) Recovering heading for visually-guided navigation. *Vision Res* 32:1177–1192. [CrossRef Medline](#)
- Kim H, DeAngelis GC (2010) Estimation of heading in the presence of moving objects: a functional role for ‘opposite’ cells in area MSTd? San Diego: Society for Neuroscience.
- Körding K (2007) Decision theory: what “should” the nervous system do? *Science* 318:606–610. [CrossRef Medline](#)
- Layton OW, Mingolla E, Browning NA (2012) A motion pooling model of visually guided navigation explains human behavior in the presence of independently moving objects. *J Vis* 12:pii20. [CrossRef Medline](#)
- Logan DJ, Duffy CJ (2006) Cortical area MSTd combines visual cues to represent 3-D self-movement. *Cereb Cortex* 16:1494–1507. [CrossRef Medline](#)
- MacNeilage PR, Zhang Z, DeAngelis GC, Angelaki DE (2012) Vestibular facilitation of optic flow parsing. *PLoS One* 7:e40264. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Mapstone M, Logan D, Duffy CJ (2006) Cue integration for the perception and control of self-movement in ageing and Alzheimer’s disease. *Brain* 129:2931–2944. [CrossRef Medline](#)
- Matsumiya K, Ando H (2009) World-centered perception of 3D object motion during visually guided self-motion. *J Vis* 9:11–13. [CrossRef Medline](#)
- Page WK, Duffy CJ (2003) Heading representation in MST: sensory interactions and population encoding. *J Neurophysiol* 89:1994–2013. [CrossRef Medline](#)
- Prsa M, Gale S, Blanke O (2012) Self-motion leads to mandatory cue fusion across sensory modalities. *J Neurophysiol* 108:2282–2291. [CrossRef Medline](#)
- Prsa M, Jimenez-Rezende D, Blanke O (2015) Inference of perceptual priors from path dynamics of passive self-motion. *J Neurophysiol* 113:1400–1413. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Royden CS (2002) Computing heading in the presence of moving objects: a model that uses motion-opponent operators. *Vision Res* 42:3043–3058. [CrossRef Medline](#)
- Royden CS, Hildreth EC (1996) Human heading judgments in the presence of moving objects. *Percept Psychophys* 58:836–856. [CrossRef Medline](#)
- Royden CS, Hildreth EC (1999) Differential effects of shared attention on perception of heading and 3-D object motion. *Percept Psychophys* 61:120–133. [CrossRef Medline](#)
- Sasaki R, Angelaki DE, DeAngelis GC (2012) Role of vestibular signals in dissociating self-motion and object motion in macaque area MSTd. New Orleans: Society for Neuroscience.
- Takahashi K, Gu Y, May PJ, Newlands SD, DeAngelis GC, Angelaki DE (2007) Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. *J Neurosci* 27:9742–9756. [CrossRef Medline](#)
- Tanaka K, Saito H (1989) Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 62:626–641. [Medline](#)
- Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J Neurosci* 6:134–144. [Medline](#)

- Tanaka K, Fukada Y, Saito HA (1989) Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 62:642–656. [Medline](#)
- Telford L, Howard IP (1996) Role of optical flow field asymmetry in the perception of heading during linear motion. *Percept Psychophys* 58:283–288. [CrossRef Medline](#)
- Telford L, Howard IP, Ohmi M (1995) Heading judgments during active and passive self-motion. *Exp Brain Res* 104:502–510. [Medline](#)
- Warren PA, Rushton SK (2007) Perception of object trajectory: parsing retinal motion into self and object movement components. *J Vis* 7:1–11. [CrossRef Medline](#)
- Warren PA, Rushton SK (2009a) Optic flow processing for the assessment of object movement during ego movement. *Curr Biol* 19:1555–1560. [CrossRef Medline](#)
- Warren PA, Rushton SK (2009b) Perception of scene-relative object movement: optic flow parsing and the contribution of monocular depth cues. *Vision Res* 49:1406–1419. [CrossRef Medline](#)
- Warren WH Jr, Morris MW, Kalish M (1988) Perception of translational heading from optical flow. *J Exp Psychol Hum Percept Perform* 14:646–660. [CrossRef Medline](#)
- Warren WH Jr, Saunders JA (1995) Perceiving heading in the presence of moving objects. *Perception* 24:315–331. [CrossRef Medline](#)
- Wichmann FA, Hill NJ (2001a) The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys* 63:1293–1313. [CrossRef Medline](#)
- Wichmann FA, Hill NJ (2001b) The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept Psychophys* 63:1314–1329. [CrossRef Medline](#)
- Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. *Science* 269:1880–1882. [CrossRef Medline](#)