Multisensory Self-Motion Compensation During Object Trajectory Judgments

Kalpana Dokka[1], Paul R. MacNeilage[2], Gregory C. DeAngelis[3] and Dora E. Angelaki[1]

[1]Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA, [2]German Center for Vertigo and Balance Disorders, University Hospital of Munich, Munich, Germany and [3]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA

Address correspondence to: Dora E. Angelaki, Department of Neuroscience, Room S740, MS: BCM295, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. Email: angelaki@bcm.edu

Kalpana Dokka and Paul R. MacNeilage contributed equally to this work.

Judging object trajectory during self-motion is a fundamental ability for mobile organisms interacting with their environment. This fundamental ability requires the nervous system to compensate for the visual consequences of self-motion in order to make accurate judgments, but the mechanisms of this compensation are poorly understood. We comprehensively examined both the accuracy and precision of observers' ability to judge object trajectory in the world when self-motion was defined by vestibular, visual, or combined visual–vestibular cues. Without decision feedback, subjects demonstrated no compensation for self-motion that was defined solely by vestibular cues, partial compensation (47%) for visually defined self-motion, and significantly greater compensation (58%) during combined visual–vestibular self-motion. With decision feedback, subjects learned to accurately judge object trajectory in the world, and this generalized to novel self-motion speeds. Across conditions, greater compensation for self-motion was associated with decreased precision of object trajectory judgments, indicating that self-motion compensation comes at the cost of reduced discriminability. Our findings suggest that the brain can flexibly represent object trajectory relative to either the observer or the world, but a world-centered representation comes at the cost of decreased precision due to the inclusion of noisy self-motion signals.

Keywords: flow parsing, object motion, optic flow, self-motion, vestibular

Introduction

We exist in a dynamic environment in which objects constantly move around us, so accurate judgment of object trajectories is essential. Assuming that the eyes and head remain stationary relative to the body, object trajectory relative to the observer can be recovered directly from retinal image motion associated with the object. We refer to this as observer-relative object motion because object motion is represented in observer-centered (i.e., egocentric) coordinates. If the observer is stationary, this estimate also conveys accurate information about how the object is moving in world (i.e., allocentric) coordinates, which we refer to as world-relative object motion. Even when the observer is moving, observer-relative object motion is sufficient for simple interception and avoidance tasks, such as catching a ball while running or avoiding collision with another moving car on the highway.

However, in many situations, we need to compute world-relative object motion while we ourselves are moving. Consider a parent walking a few steps behind a toddler in a crowded pedestrian environment. The moving parent must judge the trajectories of the child, other pedestrians, and street-fixed obstacles to protect the child from collision. To do this, the parent must compensate for self-motion to accurately judge object trajectories in world coordinates because retinal image motion is determined by the combination of object motion in the world and self-motion (Fig. 1A,B). If self-motion is not properly accounted for, then judgments of world-relative object motion will be biased. While compensating for self-motion may be beneficial for some tasks, we hypothesize that this compensation comes at the cost of decreased precision in judging object motion, because compensation should add noise associated with estimation of self-motion.

Previous studies most often report partial compensation for self-motion during judgments of object motion, but the relationship between accuracy and precision of these judgments has not been evaluated systematically. One group of studies has evaluated compensation by measuring perception of object speed. Results often show that the object had to move in the same direction as the observer to appear stationary in the world, and thus, self-motion compensation was incomplete. This is true whether the cue to self-motion was nonvisual-only (Gogel and Tietz 1977; Gogel 1982; Gogel and Tietz 1992; Mesland et al. 1996; Wexler 2003; Dyde and Harris 2008; Dupin and Wexler 2013) or visual-only (Matsumiya and Ando 2009; Dupin and Wexler 2013). During combined visual and nonvisual self-motion, more complete compensation has been reported, especially when self-motion is actively generated (Dyde and Harris 2008; Dupin and Wexler 2013). Misperception of object speed in the world has been attributed to misperception of object distance (Gogel and Tietz 1977; Gogel 1981, 1982; Gogel and Tietz 1992), misperception of self-motion velocity (Dyde and Harris 2008; Dupin and Wexler 2013), or a default tendency to judge object motion in retinal coordinates (Shebilske and Proffitt 1981).

Another group of studies has evaluated compensation by measuring perception of object trajectory. In these studies, self-motion is most often simulated based on visual cues only (but see Fajen and Matthis 2013). These studies have concluded that object direction judgments are biased in a manner consistent with the operation of a "flow-parsing" mechanism that subtracts the global component of optic flow associated with self-motion from the retinal image to estimate object motion (Gray et al. 2004; Warren and Rushton 2007, 2008; Matsumiya and Ando 2009; Warren and Rushton 2009a, 2009b; Fajen and Matthis 2013). Unfortunately, results from most of these experiments do not allow one to directly assess the degree of compensation (but see Matsumiya and Ando 2009), and thus, the nature of the representation of object trajectory (i.e., world-relative vs. observer-relative) remains unclear.
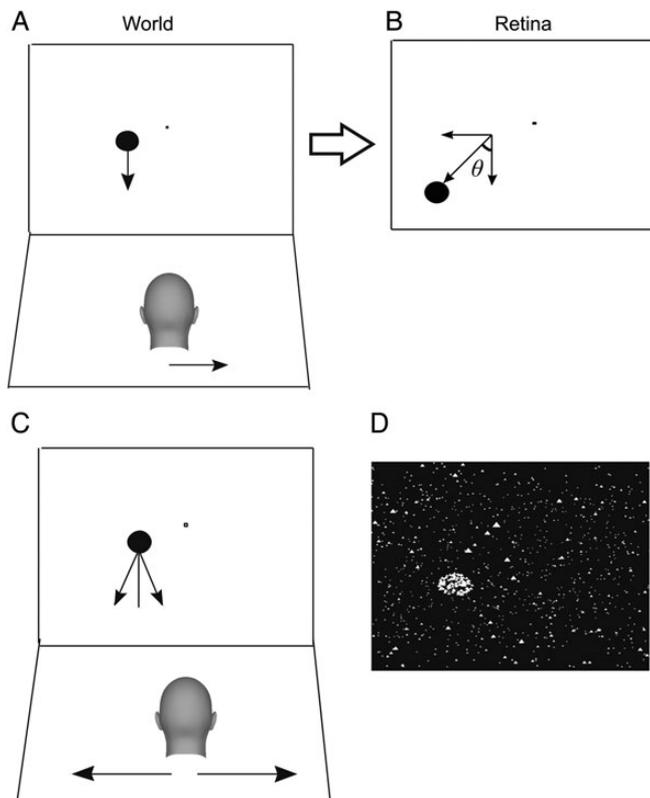
Figure 1. Schematic illustration of interactions between object and observer motion and the experimental protocol. (*A*) An object (black circle) located to the left of the observer-fixed fixation point moves downward in the world, while the subject is translated rightward. (*B*) This stimulus results in a retinal motion vector that can be computed by the vector sum of components associated with object movement (speed V_o) and observer movement (speed V_s). The direction of retinal motion (θ) is $\tan^{-1}$ (V_s/V_o). (*C*) The task was a rightward/leftward object trajectory discrimination (around a purely downward movement) in the world, while subjects also experienced rightward or leftward self-motion at 1 of 3 different velocities. (*D*) Screenshot showing the object within a starfield.

While the focus of our study is on compensating for translational self-motion during perception of object trajectory, there are a number of related studies investigating compensation associated with eye movements. For example, studies that have examined the influence of smooth pursuit, saccadic, and vergence eye movements on judging the memorized location of objects have also reported biases both during and after eye movements (Brenner et al. 2001; Blohm et al. 2003, 2005, 2006; Van Pelt and Medendorp 2008; Daye et al. 2010; Dessing et al. 2011). In particular, Van Pelt and Medendorp (2008) found that reaching movements to memorized object locations after intervening vergence movements were encoded in retinal coordinates. However, most of these studies examined localization of stationary targets and did not examine the representation of object motion in specific coordinate frames.

For self-motion compensation in most real-world situations, both visual and vestibular cues are available simultaneously and provide complementary information (Gu et al. 2007, 2008; Fetsch et al. 2009). However, to our knowledge, no previous study has examined whether the combination of visual and vestibular self-motion cues during passive motion allows for more accurate judgments of world-relative object trajectory, compared with when self-motion is indicated by only visual or vestibular cues. This study unifies the divergent research threads described above, using a single paradigm to examine

both the accuracy and precision of object trajectory judgments in world coordinates when self-motion information is provided by visual cues alone, vestibular cues alone, or combined visual–vestibular cues.

Our findings suggest that the brain can represent object trajectory in world coordinates, and that this representation is flexible. In the absence of feedback, greater evidence of self-motion (e.g., during combined visual–vestibular stimulation) increases the accuracy of world-relative judgments of object trajectory, but reduces the precision of these judgments. This tradeoff between accuracy and precision is consistent with the hypothesis that compensation adds noise associated with internal signals regarding self-motion. With feedback training, subjects learn to greatly increase the accuracy of their world-relative object judgments, but this also comes at the cost of reduced discriminability. This precision–accuracy tradeoff may reflect the weighted sum of a noisy allocentric estimate of object trajectory and a less-noisy egocentric estimate, with weights determined by the strength of evidence favoring the alternative interpretations of retinal image motion. This behavior appears to be broadly consistent with recent theories of causal inference (Kording et al. 2007).

Materials and Methods

Subjects
Nineteen healthy young adults (18–39 years) participated in the study. Subjects had no history of musculoskeletal or neurological disorders and had normal or corrected-to-normal vision. Subjects were informed of the experimental procedures, and informed written consent was obtained as per the guidelines of the Institutional Review Board.

Experimental Apparatus
Subjects were seated in a padded racing car seat mounted on a 6 degree-of-freedom motion platform (Moog 6DOF2000E). A 3-chip DLP projector (Galaxy 6; Barco, Inc.) was also mounted on the motion platform behind the subject. The projector front projected visual images onto a large projection screen (149 cm × 127 cm) via a mirror mounted above the subjects' head. The projection screen was located approximately 64 cm in front of the subject's eyes and remained fixed with respect to the subject's head. A 5-point safety harness held subjects' bodies securely in place and a custom-fitted thermoplastic mesh mask secured the head against a cushioned head mount thereby immobilizing the head relative to the chair. Subjects wore active stereo shutter glasses (CrystalEyes 3, RealD, Inc.) to provide stereoscopic depth cues. With the shutter glasses, subjects' effective field of view was approximately 90° × 70°. In some subjects, the eye position was recorded for both eyes at 600 Hz via a video-based eye-tracking system (ISCAN) attached to the stereo glasses. Sounds from the platform were masked by playing white noise through headphones. Behavioral responses were collected using a button box.

Stimuli and Procedure
During each trial, subjects viewed the movement of a circular object (diameter 15°) that moved downward with a leftward or rightward component relative to earth-vertical (Fig. 1*C*). Subjects were instructed to judge the direction of object motion in the world. That is, they were asked to indicate, via a button press, whether the object moved rightward or leftward in the world. These instructions assume that subjects perceive the virtual world rendered on the visual display as spatially congruent with the real world. Although this is an assumption, nothing in our debriefing of subjects suggested otherwise. In each trial (1 s duration), subjects fixated a centrally-located, head-fixed fixation point. The left/right deviation of the object with respect to earth-vertical was varied from trial to trial using a 1-up/1-down staircase procedure. The maximum angle of deviation from vertical was ±60°. In trials with self-motion, subjects experienced either rightward or leftward translation in

the fronto-parallel plane (interaural axis). Self-motion was chosen to be approximately orthogonal to the direction of object motion, thus allowing us to probe how self-motion signals impact perception of object trajectory rather than just object speed.

Self-motion was either visually simulated (by optic flow) and/or provided by inertial motion of the platform. Self-motion and object motion were synchronized and both followed Gaussian velocity profiles in the horizontal direction; object motion followed a similar Gaussian velocity profile in the vertical direction. The Gaussian velocity profile had a standard deviation = 1/6th s. We elected to have the motion trajectories follow a Gaussian velocity profile because: (1) It is a smooth, transient, and natural stimulus, (2) it evokes robust visual and vestibular responses in cortical multisensory neurons, such as those in the dorsal medial superior temporal (MSTd) and ventral intraparietal (VIP) (Gu et al. 2006; Chen et al. 2011), (3) it results in near-optimal multisensory integration, both at the level of behavior (Gu et al. 2008; Fetsch et al. 2009) and in neuronal activity (Morgan et al. 2008; Fetsch et al. 2011).

Subjects experienced rightward or leftward self-motion at 3 peak velocities of 12, 22, and 32 cm/s (peak displacements: 6, 11, and 16 cm, respectively; peak accelerations: 0.37, 0.68, and 0.99 m/s^2, respectively). Peak retinal image velocity of the object (V_{obj}) was always 32°/s. The horizontal and vertical velocity components of the object were equal to $V_{obj} \sin \theta$ and $V_{obj} \cos \theta$, respectively, where θ indicates the angular deviation of the object from earth-vertical.

The background in the visual scene (Fig. 1D) consisted of a 3-dimensional starfield (170 cm × 170 cm × 100 cm) that was composed of randomly placed triangles having a base and height of 1.6 cm and a density of 0.0005 triangles/cm^3 (~1445 triangles per video frame). The moving object was also composed of randomly placed triangles, with a density of 0.58 triangles/cm^2 (~116 triangles per frame). The starfield was centered around the screen, such that it began 50 cm behind the screen and extended up to 50 cm in front of the screen. The object was rendered at the same depth as the screen and was thus located in the center of the starfield. The triangles composing both the background and the object were opaque such that a nearer triangle could occlude part or all of a farther triangle. However, both the background and the object were otherwise transparent, such that background triangles were visible through the object or vice versa. Apart from the triangle-specific occlusions described above, there were no accretion/deletion cues at the boundary of the object. Motion coherence of both starfield and object was set to 100%, and the object was easily detectable in the presence of the starfield. The object was initially located to the left of the fixation point at a distance of 25–35°, and the initial horizontal position of the object was varied randomly within this range from trial to trial.

Main Experiment

Subjects reported the direction of object motion in the world during lateral self-motion. There were 4 self-motion conditions: (1) Object motion only, no self-motion (i.e., motion platform stationary; background starfield not present) (Obj); (2) object motion with passive vestibular self-motion (i.e., motion of the platform only; background starfield not present) (Obj + Vest); (3) object motion with visually simulated self-motion (i.e., motion of the background starfield only; motion platform stationary) (Obj + Vis); and (4) object motion with combined visual–vestibular self-motion, in the form of synchronized optic flow and motion of the platform (Obj + Com). In trials with combined visual–vestibular self-motion, the background triangles moved on the screen in a direction opposite to that of inertial motion of the platform. That is, during rightward self-motion, background triangles moved to the left on the display and vice versa for leftward self-motion. We refer to the last 3 stimulus configurations above as the vestibular, visual, and combined conditions, respectively. While it is true that there are nonvisual cues other than vestibular cues during passive self-motion, prior research using very similar stimulus parameters has demonstrated that it is primarily vestibular signals that limit sensitivity for self-motion judgments (Gu et al. 2007; Takahashi et al. 2007).

Twelve subjects (S1–S12) performed the object trajectory discrimination task under all the 4 self-motion conditions. Subjects experienced rightward and leftward self-motion in separate sessions. Within

each session (with either rightward or leftward self-motion), subjects experienced 10 interleaved stimulus conditions: Obj, Obj + Vest at 3 self-motion velocities, Obj + Vis at 3 self-motion velocities, and Obj + Com at 3 self-motion velocities. In each session, there were 40 trials for each distinct stimulus condition and subjects participated in 3 sessions each for rightward and leftward self-motion directions. Thus, each subject performed a total of 2400 trials (40 trials per condition × 10 conditions per session × 3 sessions × 2 self-motion directions). In this main experiment, subjects did not receive feedback about their task performance. In addition, for these subjects, binocular left and right eye positions were recorded at 600 Hz using a video-based, eye-tracking system (Dokka et al. 2011).

At the end of trials with platform movement, the platform returned to a central position. Return of the platform to the central position could provide information about the speed, direction, or distance of self-motion in the preceding trial, but it could not provide feedback about the object trajectory in world coordinates because object direction was manipulated independently of platform motion. Importantly, subjects indicated their response in each trial before the platform returned to its central position, so any feedback received during the return motion of the platform cannot influence subjects' response on that trial. Although the return motion of the platform could conceivably help subjects to gauge their self-motion (while the object is not present), self-motion speeds and directions were presented in a random order. Thus, any feedback received from the return motion of the platform would be unlikely to improve subjects' sense of self-motion on the next trial. Taken together, these considerations make it extremely unlikely that the return motion of the platform provided a meaningful cue to facilitate object trajectory perception in the world.

Feedback Training Experiments

To examine whether feedback training improves the accuracy of world-relative object trajectory judgments, another set of experiments was performed following a training session in which each subject received audio feedback regarding the correctness of their decisions. Feedback training was always conducted in the vestibular self-motion condition (Obj + Vest). Vestibular self-motion was chosen as the training condition for 2 reasons. First, before feedback training, the largest biases in perceived object trajectory were observed with vestibular self-motion (see Results—Main experiment). Training with this condition allowed us to maximize the potential effects of feedback training. Second, training only with vestibular self-motion allowed us to examine whether there was a cross-modal transfer of training effects to the visual and combined self-motion conditions, which did not receive feedback training. Feedback training was implemented using different audio sounds. One audio tone indicated that subjects correctly judged object trajectory in the world, while a different tone indicated that their judgment was incorrect. Subjects were given a brief rest after the feedback training session; subsequently, they participated in a test session in which there was no feedback.

In these feedback experiments, rightward and leftward self-motion directions were interleaved in both training and test sessions. As interleaving the self-motion directions with stimulus modality and speed greatly increased the duration of each experimental session, it was not possible to test all self-motion conditions within a single session. Therefore, vestibular self-motion (Obj + Vest) was tested in some sessions, whereas both visual (Obj + Vis) and combined (Obj + Com) self-motion conditions were interleaved in other test sessions. Testing visual and combined conditions in separate sessions allowed us to examine whether the effects of vestibular feedback training generalized to the visual and combined conditions.

The feedback training that preceded each test session consisted of 90 trials (15 trials × 3 velocities of vestibular self-motion × 2 directions). Test sessions with vestibular self-motion consisted of 240 trials (40 trials × 3 self-motion velocities × 2 directions). Test sessions with visual and combined self-motion consisted of 480 trials (40 trials × 3 self-motion velocities × 2 directions × 2 conditions). Each type of test session was repeated 3 times. Thus, for vestibular test sessions, there were a total of 270 training trials and 720 test trials. For visual and

combined test sessions, there were a total of 270 training trials and 1440 test trials.

Vestibular test sessions, following feedback training, were conducted with 4 previously tested subjects (S6, S7, S8, and S12) and 5 new subjects (S13, S14, S15, S18, and S19). For the 5 new subjects, baseline data for the vestibular condition (Obj + Vest) were first collected without feedback. Only then did these subjects participate in the experiments with feedback training. To investigate whether the effects of feedback training generalized to novel stimuli, 4 of these subjects (S7, S12, S13, and S18) experienced novel self-motion velocities in the test session. That is, in addition to the 3 self-motion velocities for which they received feedback in the training session (12, 22, and 32 cm/s), these subjects also experienced self-motion velocities of 16 and 26 cm/s in the test session. These 4 subjects completed 400 trials per test session (40 trials × 5 self-motion velocities × 2 directions) and 1200 trials total (400 trials × 3 repetitions).

The visual and combined test sessions, following feedback training, were conducted with 5 previously tested subjects (S2, S3, S6, S8, and S12) and 5 of the new subjects (S15, S16, S17, S18, and S19). Here too, data were first collected without feedback in the 5 new subjects. Only then did they participate in experiments with feedback training.

Data Analysis

For each condition in each session, a trajectory estimation bias is quantified by measuring the point of subjective equality, which indicates the object trajectory angle perceived to be earth-vertical. This bias in the perceived direction of object motion was calculated as an average of staircase reversals (excluding the first 4 reversals and considering the largest even number of reversals). The mean bias for each self-motion condition was obtained by averaging the biases across all 3 sessions. In addition, bias was also calculated by fitting psychometric (cumulative Gaussian) functions to the data pooled across sessions (Wichmann and Hill 2001a; 2001b). The mean of the cumulative Gaussian fit provides an additional measure of bias. Bias values calculated these 2 different ways were very similar (Supplementary Fig. 1A–D); therefore, we only report bias values calculated according to the staircase method. The object trajectory discrimination threshold, or just noticeable difference, quantifies the precision of object motion estimates. We define threshold to be equal to the standard deviation of the cumulative Gaussian fit to the psychometric function. The coefficient of determination values (R^2) were calculated (as unity minus the ratio of the sum of squared error between the data and psychometric fit to the total sum of squares) to evaluate the goodness of psychometric fits (Supplementary Fig. 1E,F).

To quantify the degree to which subjects accounted for self-motion when judging object trajectory, we calculated percentage compensation as follows: % compensation = $(1 - bias_{meas}/bias_{ret}) \times 100$. Here, $bias_{meas}$ represents the empirical bias in perceived object trajectory, whereas $bias_{ret}$ represents the predicted bias if subjects judged object motion in purely retinal (observer) coordinates (i.e., no compensation for self-motion). This predicted bias is computed as the inverse tangent of the ratio of the self-motion and object motion components of retinal image speed associated with the moving object (Fig. 1B).

Slow-phase eye velocity was calculated in the following manner. Trials in which subjects blinked were characterized by abrupt step transients and were eliminated from the eye movement analysis. The horizontal position trace of the left eye was smoothed and differentiated using a Savitzky-Golay moving window filter to yield eye velocity (Savitzky and Golay 1964; Press et al. 1992). Eye velocity was successively differentiated to yield acceleration and jerk. A jerk threshold was implemented to identify and eliminate eye saccades (Wyatt 1998). In each trial, slow-phase velocity was averaged across the duration of the trial. For each stimulus condition, mean slow-phase velocity was calculated by averaging across trials.

When presenting averages, the mean, standard error of the mean (SEM), and 95% confidence interval (95% CI) are reported. For statistical analysis of perceptual biases, values obtained with leftward self-motion were multiplied by −1 to allow comparison with bias values obtained during rightward self-motion. Repeated-measures analysis of variance (ANOVA) was used to analyze data from experiments without

feedback training. Note that for the feedback experiments, visual and combined conditions were tested together, while vestibular condition was tested in different sessions, so feedback data were analyzed using a standard rather than repeated-measures ANOVA. These analyses were used to analyze bias, threshold, and % compensation data, with the factors being self-motion condition, velocity, and movement direction. Tukey–Cramer post hoc comparisons were also used; in this case, data were pooled across factors not being tested. Linear regression and paired t-tests were also performed to compare bias and threshold values across conditions.

Note that the initial position of the object was randomly varied from 25° to 35° to the left of the fixation point to discourage subjects from basing some decisions on object position. To assess the efficacy of this manipulation, we analyzed the influence of object starting position on the bias in perceived object motion before feedback training. Trials were grouped into 2 bins of starting object position: 25–30° and 30–35°. For each condition and each bin, psychometric functions were constructed to calculate the bias in perceived object motion. The biases estimated for the 2 position bins were compared using repeated-measures ANOVA, which tests the null hypothesis that bias is independent of the starting eccentricity of the object. Over the range of positions (10°) tested, we find no evidence to reject this null hypothesis (repeated-measures ANOVA, $F_{1,419} = 0.19$, $P = 0.66$).

Results

Subjects were required to judge whether an object moved downward and leftward or downward and rightward in world coordinates (Fig. 1). Four stimulus conditions were tested: (1) Object motion only (Obj: motion platform stationary; background starfield not present), (2) object motion with passive vestibular self-motion (Obj + Vest: motion of the platform only; background starfield not present), (3) object motion with visually simulated self-motion (Obj + Vis: motion of the background starfield only; motion platform stationary), and (4) object motion with combined visual–vestibular self-motion (Obj + Com: synchronized motion of the platform and background starfield).

Main Experiment

As illustrated by representative data from one of the subjects, object motion alone (without self-motion) yielded staircases that converged toward earth-vertical estimates, with close to zero bias (Fig. 2A,B, dashed-gray). Corresponding fits of the psychometric data with cumulative Gaussian functions were steep with close to zero mean (Fig. 2C,D, dashed-gray), indicating that object motion judgments were rather accurate in the absence of self-motion. In contrast, large systematic biases in object motion perception were observed during the vestibular self-motion condition (Obj + Vest; Fig. 2A–D, thick-gray), with opposite direction biases for leftward and rightward self-motion. Smaller, but still systematic, biases were also observed when self-motion was simulated using optic flow (Obj + Vis; Fig. 2A–D, thin-black). For both directions of self-motion, the object trajectory needed to have a horizontal component in the same direction as the subject's self-motion in order to be perceived as moving vertically in the world. When self-motion was specified by congruent combinations of visual and vestibular cues, the bias was further reduced (Obj + Com; Fig. 2A–D, thick-black; ANOVA, $P < 0.001$). Results are summarized for this subject in Figure 2E,F for all 3 speeds of self-motion. The larger the self-motion speed, the larger the observed bias in perceived object trajectory, and the larger the reduction in bias when both visual and vestibular cues were presented together (Obj + Com).
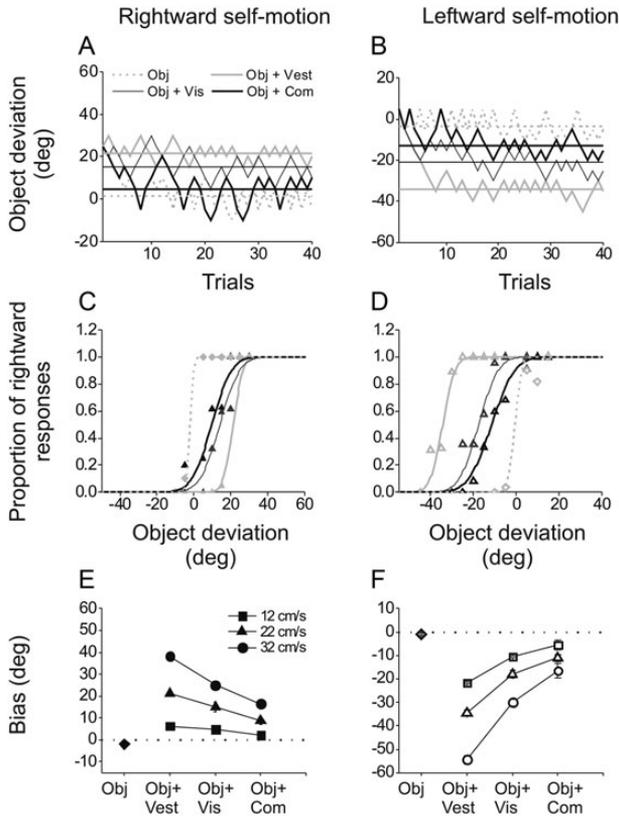
Figure 2. Data from a representative subject. (*A* and *B*) Staircase histories of a representative subject during individual sessions, shown separately for rightward and leftward self-motion (peak velocity of 22 cm/s). Data are shown for perceived object trajectory judgments in the absence of self-motion (Obj, dashed-gray), with vestibular self-motion (Obj + Vest, thick-gray), with optic flow (Obj + Vis, thin-black), and with combined visual–vestibular self-motion cues (Obj + Com, thick-black). Color- and style-coded solid horizontal lines indicate averages of staircase reversals used to compute object trajectory bias (see Materials and Methods). (*C* and *D*) Psychometric functions for this subject's cumulative data during rightward (filled symbols) and leftward (open symbols) self-motion (peak velocity of 22 cm/s), coded as in *A* and *B*. (*E* and *F*) Average bias values for this subject across sessions, during rightward and leftward self-motion, respectively. Dotted horizontal black line illustrates zero bias. Error bars indicate the SEM.

Figure 3. Average bias and thresholds across subjects. (*A* and *B*) Mean bias in the perceived object direction for rightward (filled symbols) and leftward (open symbols) self-motion, both before (*A*) and after (*B*) feedback training. Note that 12 subjects were tested in all conditions before feedback training, whereas 9 and 10 subjects were tested with vestibular and visual/combined self-motion after feedback training, respectively. Dotted horizontal black lines illustrate zero bias. The solid horizontal (corresponding to self-motion velocity of 12 cm/s), dashed (corresponding to 22 cm/s), and dash-dotted (corresponding to 32 cm/s) lines indicate the bias expected if subjects judged object trajectory in observer-centered coordinates. (*C* and *D*) Mean object trajectory discrimination thresholds during rightward (top) and leftward (bottom) self-motion before (*C*) and after (*D*) feedback. Error bars indicate SEM.

A similar pattern of results was observed for all subjects. As shown in Figure 3*A*, the mean bias in perceived object direction across subjects depended significantly on the stimulus condition (repeated-measures ANOVA; factors: Self-motion condition, self-motion speed, and self-motion direction; main effect of self-motion condition: $F_{2,627} = 407.6$, $P < 0.0001$). Post hoc comparisons show that the bias measured during the vestibular condition was significantly greater than that for the visual condition ($P < 0.001$, data pooled across self-motion velocity and direction, see Materials and Methods), which was, in turn, significantly greater than the bias measured in the combined condition ($P < 0.001$). This indicates that combined visual and vestibular self-motion cues significantly reduced the bias in perceived object motion.

If subjects did not fixate accurately during real or simulated self-motion, this could be a potential confound in the results. Although average eye velocities were small, mean eye velocity did exhibit a significant dependence on self-motion speed (repeated-measures ANOVA; factors: Stimulus condition, speed, and direction; main effect of speed: $F_{2,627} = 10.1$, $P < 0.001$) and
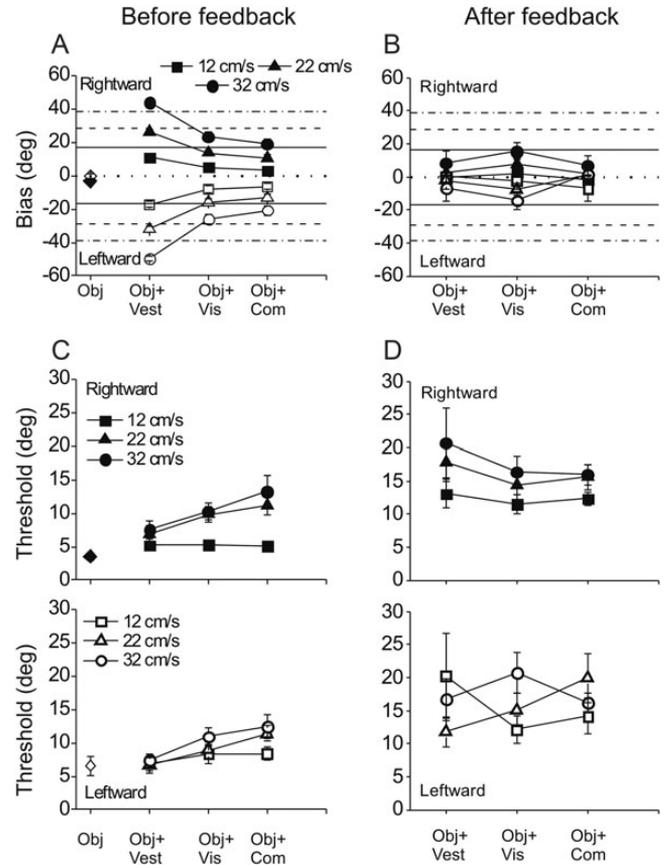
stimulus condition ($F_{2,627} = 40.1$, $P < 0.001$; Supplementary Fig. 2*A*,*B* and Supplementary Table 1). There was a significant difference in mean eye velocity at 12 cm/s when compared with 22 ($P < 0.05$) and 32 ($P < 0.05$) cm/s. Furthermore, eye velocity in the combined condition was significantly greater than that in the visual condition ($P < 0.05$), which, in turn, was significantly greater than in the vestibular condition ($P < 0.05$). Critically, however, there was no significant correlation between mean eye velocity and bias in perceived object motion (data pooled across self-motion conditions, speeds, and directions; Pearson correlation: $R = 0.06$; $P = 0.33$; Supplementary Fig. 2*C*), suggesting that differences in perceptual judgments were not a confound of eye movements.

Compared with the bias data of Figure 3*A*, an inverse dependence on self-motion conditions was observed for the precision of object motion judgments (Fig. 3*C*). That is, psychophysical thresholds for object motion discrimination tended to increase from the vestibular to the visual to the combined condition. When data were pooled across all 3 self-motion speeds, there was a significant influence of the stimulus condition on the

threshold values (repeated-measures ANOVA; factors: Stimulus condition, speed, and direction; main effect of stimulus condition: $F_{2,195} = 13.08$, $P < 0.001$). Post hoc comparisons revealed a significant difference between vestibular and visual conditions ($P = 0.03$), but the difference between visual and combined conditions was not significant ($P = 0.08$). However, this analysis revealed a significant interaction between stimulus condition and self-motion speed (repeated-measures ANOVA; factors: Stimulus condition, speed, and direction; significant interaction between stimulus condition and speed: $F_{4,195} = 8.24$, $P < 0.001$), such that the increase in thresholds from the vestibular to visual to combined conditions was greater for the 2 larger self-motion speeds. Taken together, these data show that more accurate object trajectory judgments were also less precise.

To better understand subjects' response strategy, we compared experimentally measured biases with the biases that would be predicted if subjects made judgments in observer (i.e., retinal) coordinates. The predicted bias is computed as the inverse tangent of the ratio of the self-motion component to the object motion component of retinal image speed associated with the object (Fig. 1B). The predicted biases, if judgments were made in observer coordinates, are 16.7°, 28.8°, and 38.6° for self-motion speeds of 12, 22, and 32 cm/s, respectively (horizontal solid, dashed, and dash-dotted lines in Fig. 3A). Figure 4 plots measured biases versus those predicted by an observer-relative strategy. The slope of the best-fit line quantifies the average ratio of measured-to-predicted biases for each stimulus condition. This slope was close to 1 for the vestibular condition (Fig. 4A, blue; slope = 1.11, 95% CI = 1.05–1.17), indicating that subjects tended to judge object motion in observer coordinates during self-motion in the absence of optic flow. In the visual condition, the slope was 0.60 (95% CI = 0.51–0.69), suggesting partial compensation for self-motion (Fig. 4C, green). Finally, in the combined condition, the slope of the relationship between measured and predicted biases was reduced further (slope = 0.40; 95% CI = 0.33–0.54, Fig. 4C, red), consistent with the hypothesis that compensation is most complete when both visual and vestibular cues to self-motion are available.

To evaluate how much subjects compensated for their self-motion when judging object trajectory, percent compensation was computed for each subject as 1 minus the ratio of observed-to-predicted bias, multiplied by 100 (see Materials and Methods). Zero percent compensation indicates that subjects judged object motion in observer coordinates (i.e., following an observer-relative strategy), whereas 100% compensation indicates that subjects accurately judged object motion in world coordinates. Percent compensation, averaged across self-motion speeds and movement directions, was −2.4% (95% CI = −6.8–2.1) for the vestibular condition, 46.9% (95% CI = 42.4–51.3) for the visual condition, and 57.8% (95% CI = 53.3–62.3) for the combined condition (Fig. 4E). Note that the 95% CI for the vestibular condition includes 0, suggesting that judgments were made in observer coordinates during this condition. Furthermore, the 95% CI for the combined condition did not overlap with 95% CI for the visual or vestibular conditions. Thus, the combination of vestibular and visual self-motion signals leads to a significant improvement in the ability of subjects to compensate for self-motion and to judge object trajectory in the world.

Recall that an inverse relationship appears to exist between accuracy (bias) and precision (threshold) for judging object trajectory in the world (Fig. 3A,C). To examine this precision–accuracy tradeoff in greater detail, we computed the correlation
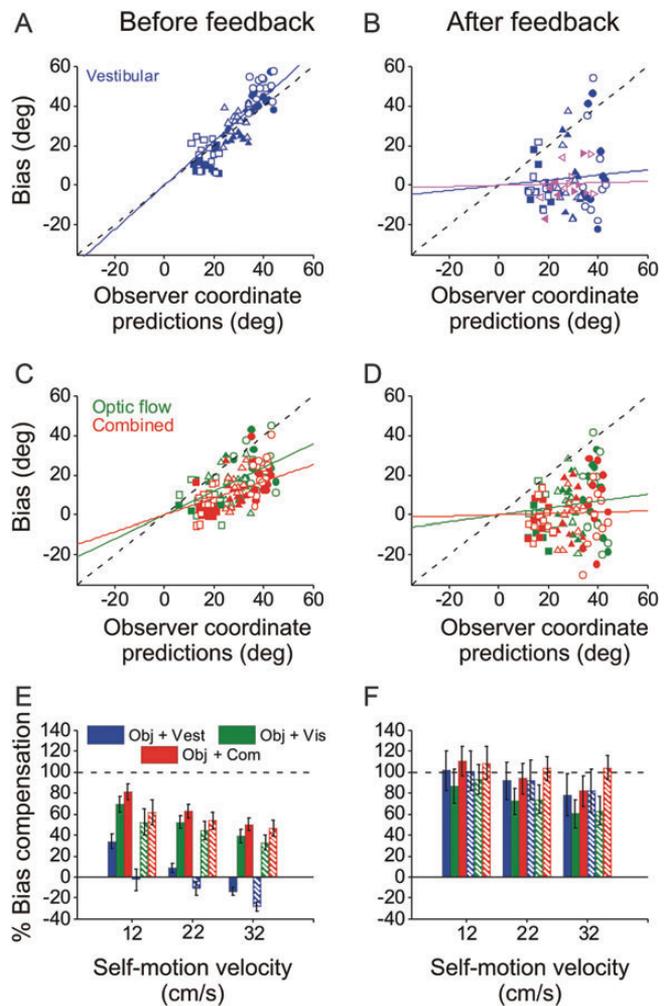

Figure 4. Comparison between bias and observer coordinate predictions. Data are shown separately for the vestibular condition (A and B), as well as the visual (green) and combined (red) conditions (C and D). Left and right columns show data obtained before and after feedback training. Magenta symbols in B represent data obtained from 2 novel self-motion velocities for which no feedback was ever given to subjects. Black dashed lines represent unity slope. Solid colored lines represent best-fitting linear models obtained from regression. Note that the slope of this line quantifies the average ratio of measured to predicted bias for each data set. Small random offsets were added to the observer coordinate predictions to allow better visualization of the data. Twelve subjects were tested in all conditions before feedback training, whereas 9 and 10 subjects were tested with vestibular and visual/combined self-motion after feedback, respectively. (E and F) Percent compensation before and after feedback. Dashed horizontal line represents perfect (100%) compensation for self-motion. Error bars indicate SEM. Filled and hatched bars correspond to rightward and leftward self-motion directions, respectively.

between threshold and percent compensation across all subjects and conditions, with self-motion speed as a covariate. Indeed, percent compensation and threshold values were significantly correlated (analysis of covariance; $R = 0.25$, $P = 0.01$), even after accounting for the effect of self-motion velocity. Thus, the process of compensation for self-motion appears to add noise to object motion estimates (see Discussion).

In addition, the percent compensation metric shows significant dependencies on self-motion speed and direction (Fig. 4E). Percent compensation was significantly greater for slower speeds of self-motion (repeated-measures ANOVA: Factors: Stimulus condition, speed, and direction; main effect of speed: $F_{2,627} = 57.33$, $P < 0.001$), suggesting that subjects have more

difficulty accounting for self-motion at faster speeds. Percent compensation was also greater for rightward than leftward self-motion (repeated-measures ANOVA: Factors: Stimulus condition, speed, and direction; main effect of direction: $F_{1,627} = 97.02$, $P < 0.001$).

Note that, for the fastest speeds in the vestibular condition, negative percent compensation values were sometimes observed (Fig. 4E, blue negative-going bars), indicating that the observed bias was larger than the bias predicted by the retinal strategy. This means that, for object motion perceived to be earth-vertical, the lateral component of retinal image motion associated with the object was in the same direction as the observer's self-motion. This cannot be explained by mis-estimation of self-motion or object distance, because all points nearer than infinity will generate retinal image motion opposite to the observer's self-motion, with speed inversely proportional to distance. Nor can this effect be explained by uncontrolled eye movements (Supplementary Fig. 2C). One possibility is that subjects may have incorrectly perceived relative motion between the fixation point and themselves. Such an additional motion percept could potentially lead to negative compensation values.

Feedback Training Experiment

The minimal compensation for self-motion that occurred in the vestibular condition suggests that subjects simply reported object motion in observer coordinates. This could reflect a lack of neural mechanisms to account for vestibular self-motion when judging moving objects. Alternatively, it could simply reflect the fact that subjects did not know the outcome of their decisions and, therefore, did not realize that their default strategy was insufficient under these laboratory conditions. We hypothesized that subjects could learn to account for self-motion and make world-relative judgments of object motion if they received feedback regarding their decisions. Thus, a feedback training experiment was conducted to see whether subjects could overcome their biases and report object trajectory in world coordinates. Before each test session, subjects completed a training session with vestibular self-motion (Obj + Vest) in which they were informed whether their answers were correct or not (see Materials and Methods for details).

There were marked changes in subjects' perceptual judgments after feedback training. Most importantly, there was a dramatic reduction in observed biases (Fig. 3B), not only in the vestibular condition, for which feedback was provided in the training session, but also in the visual and combined conditions, for which no feedback was ever given (paired t-tests comparing bias before and after feedback for vestibular, visual, and combined conditions; $P < 0.001$; Supplementary Fig. 3A,B). Interestingly, there was still an effect of self-motion condition on the bias after feedback training (ANOVA; factors: Self-motion condition, self-motion speed, and self-motion direction; main effect of self-motion condition: $F_{2,512} = 14.9$, $P < 0.001$). Biases in both vestibular ($P < 0.05$) and combined ($P < 0.05$) conditions were significantly reduced relative to the visual condition. However, biases in the vestibular and combined conditions were not significantly different from one another.

After feedback training, observed biases were no longer similar to those predicted by an observer-relative strategy. The slope of the relationship between observed and predicted biases decreased from 1.11 (Fig. 4A) to 0.14 (95% CI = −0.016–0.29, Fig. 4B) for the vestibular condition, from 0.60 (Fig. 4C,

green) to 0.17 (95% CI = 0.050–0.30, Fig. 4D, green) for the visual condition, and from 0.40 (Fig. 4C, red) to 0.032 (95% CI = −0.075–0.14, Fig. 4D, red) for the combined condition. Both vestibular and combined slopes were not significantly different from zero, suggesting that subjects learned to judge object motion in world coordinates. Whereas most subjects improved their performance, feedback training did not lead to decreased bias in all subjects (Fig. 4B,D, data points close to the unity slope line; see also Supplementary Fig. 4). It appears that some subjects did not learn to make use of the feedback in the vestibular condition.

The improved accuracy after feedback training is also reflected in percent compensation values (Fig. 4F). Mean values across self-motion velocities were 94.3% (95% CI = 85.4—103.9), 70.0% (95% CI = 63.5–76.6), and 97.8% (95% CI = 91.2–104.4) for the vestibular, visual, and combined conditions, respectively. The 95% CI includes 100% for both vestibular and combined conditions, thus further indicating that subjects successfully judged object trajectory in world coordinates. The lower average percent compensation in the visual condition presumably stems from the fact that feedback training was only provided in the vestibular condition, and thus, it had the greatest effect on test sessions involving the vestibular and combined conditions, for which inertial motion cues were present. Importantly, however, percent compensation values were significantly greater for the visual condition after feedback training than before training (ANOVA; factors: Feedback status (before/after), self-motion direction, and self-motion speed; main effect of feedback status: $F_{1,346} = 110.75$, $P < 0.001$). Thus, there was a significant transfer of the feedback training effect from the vestibular training condition to the visual test condition.

We again examined the dependencies of percent compensation on self-motion direction and speed. After feedback training, the asymmetry between rightward and leftward self-motion observed before feedback was no longer present (Fig. 4F; ANOVA; factors: Stimulus condition, speed, and direction; main effect of direction: $F_{1,512} = 2.60$, $P = 0.11$). That is, contrary to the result observed before feedback training, percent compensation measured with rightward self-motion was not significantly different from that measured with leftward self-motion. However, similar to before feedback training, percent compensation exhibited a significant dependence on self-motion velocity (Fig. 4F; ANOVA; factors: Stimulus condition, speed, and direction; main effect of speed: $F_{2,512} = 10.7$, $P < 0.0001$). Note that this comparison does not reveal whether training induced greater improvements in accuracy at slow self-motion compared with fast velocities. To examine this, we computed the change in percent compensation due to feedback training for all self-motion velocities. There was no significant dependence of the change in percent compensation on self-motion velocity, suggesting that feedback training induced similar improvements in accuracy for all self-motion velocities (ANOVA; factors: Stimulus condition, speed, and direction; main effect of speed: $F_{2,164} = 0.11$, $P = 0.89$).

The notion that feedback training helped subjects to learn to judge object trajectory in world coordinates is strengthened by the observation that feedback training effects generalized to novel self-motion velocities. Some subjects (see Materials and Methods) were tested with 2 novel self-motion speeds (i.e., 16 and 26 cm/s) for which they did not receive any feedback training. As seen in Figure 4B, the slope between observed and predicted bias for the novel velocities (0.044, 95% CI = −0.12–0.21,

magenta) was similar to that for the trained velocities (slope = 0.041, 95% CI = −0.04–0.12), and neither was significantly different from zero. Likewise, average percent compensation values for the novel (106.6%, 95% CI = 94.1–119.0) and trained velocities (98.5%, 95% CI = 90.0–106.91) were not significantly different from each other (t-test, t = −0.65, P = 0.52), and neither was significantly different from 100%. Thus, subjects appeared to treat the novel self-motion velocities in the test sessions identically to the velocities that received feedback training.

The decrease in bias (Fig. 3B; Supplementary Fig. 3A,B) and increase in percent compensation after feedback training (Fig. 4E,F) were accompanied by an overall significant increase in discrimination thresholds for all stimulus conditions (Fig. 3D; Supplementary Fig. 3C,D; paired t-test, P < 0.01). This observation supports the notion that feedback training increases percent compensation by allowing subjects to effectively incorporate self-motion information, and this leads to higher discrimination thresholds because noise from the self-motion estimate is incorporated into object motion judgments.

Discussion

To estimate object motion relative to the world, moving observers must compensate for their self-motion to judge object trajectory in the world (Wallach 1987; Gogel 1990; Dupin and Wexler 2013; Fajen and Matthis 2013). We performed the first systematic analysis of both the accuracy and precision of observers' object trajectory judgments in visual-only, vestibular-only, and combined visual–vestibular self-motion conditions. In the vestibular condition, without decision feedback, subjects judged object trajectory in observer coordinates, and this may have been facilitated by the sparse, observer-fixed visual reference frame comprised of the fixation point and the display boundary, as discussed further below. In the visual condition, subjects demonstrated partial compensation (47%), perhaps reflecting the competing influences of the observer-fixed cues and the world-fixed reference frame indicated by background motion. Even greater compensation (58%) was observed in the combined condition, suggesting that vestibular signals influence the estimate of self-motion derived from optic flow. Object trajectory judgments were more precise in the vestibular condition than in the visual and combined conditions, most likely because compensation incorporates noise associated with the self-motion estimate. Results of the feedback training experiment demonstrate that subjects can learn to compensate for self-motion (70–98% compensation), but this compensation also incurs a corresponding loss in precision of object motion judgments. Taken together, our findings suggest that the brain can represent object motion in allocentric coordinates, but these representations are inherently less precise due to inclusion of noise associated with estimates of self-motion.

Judging Object Motion with Nonvisual Self-Motion Cues Only

Prior research has demonstrated imperfect compensation for self-motion when subjects judge object motion in the absence of visual self-motion cues (Swanston and Wade 1988; Wexler 2003; Dyde and Harris 2008; Dupin and Wexler 2013), similar to our results. In most of these studies, subjects moved actively in a completely dark environment while viewing a single illuminated object. Under such circumstances, observers often misjudge object speed; the object must move with the observer in order to appear stationary in the world. This outcome has been interpreted in different ways. It has been shown that biases in judging object speed correlate with perceived distance of the object (Gogel 1982; Gogel and Tietz 1992). This makes sense because, for a fixed observer translation, the image of a near object will move farther and faster over the retina than that of a far object, a phenomenon known as motion parallax. Thus, the appropriate compensation for self-motion induced image motion necessarily depends on object distance.

However, misperception of object distance alone is unlikely to explain our results. As shown in Figure 4A, results of our vestibular condition are consistent with subjects judging object motion in retinal or screen coordinates, despite explicit instructions to judge object motion in world coordinates. The object had to move laterally at the same speed and direction as the observer in order for subjects to perceive earth-vertical motion. If this effect were attributable mainly to misperception of object distance, these judgments could only be explained by a perceived object distance of infinity (e.g., similar to looking at the moon from a moving car). Our stimulus was rendered on a display screen at a distance of 64 cm, with salient depth cues such as binocular disparity and accommodation signaling near depth. It is therefore highly unlikely that such a stimulus configuration led to infinite perceived distance of the object.

Alternatively, incomplete compensation may result from misperception of self-motion (Dyde and Harris 2008; Dupin and Wexler 2013). The image of a stationary object will move faster over the retina when self-motion itself is faster. Thus, error in estimating self-motion will lead to under- or over-compensation and result in erroneous object motion judgments. However, this is also unlikely to explain our findings in the vestibular condition, because our pattern of results is consistent only with a percept of zero self-motion.

We suggest instead that the observed biases in the vestibular condition may reflect a default tendency for observers to judge object motion in observer coordinates. Such a tendency has previously been proposed (Shebilske and Proffitt 1981) and can account well for our results. Indeed, as pointed out previously (Wexler 2003; Dupin and Wexler 2013), the degree of compensation in such tasks will lie somewhere along the continuum between that consistent with pure egocentric (observer-centered) and allocentric (world-centered) reference frames. Where the data fall along this continuum likely depends on experimental factors that shape the frame of reference that subjects apply to the task. For example, in our apparatus, the entire display screen and fixation target moved with the subject's head; thus, visible luminance boundaries at the edge of the display, along with the fixation target, provided observer-fixed reference points against which object motion could be judged. These stimulus features likely contributed to the tendency for subjects to judge object motion in observer coordinates in the vestibular condition (without feedback). Such observer-fixed references were absent or substantially less prominent in prior experiments (Dyde and Harris 2008; Matsumiya and Ando 2009; Dupin and Wexler 2013).

Most relevant to the current research are 2 prior studies (Wexler 2003; Dyde and Harris 2008), in which subjects were asked to judge object speed during passive self-motion in the absence of visual self-motion cues. Unlike our vestibular

condition, experiments were conducted in nearly complete darkness, such that the only visible feature was the object to be judged. Consequently, eye movements were not controlled, and subjects tracked the object with their eyes. Both of these studies report approximately 40% compensation, whereas we observed approximately 0% compensation. This difference could be explained by the additional visible display features in our experiment or by the absence of tracking eye movements in our study.

In both of these previous studies, compensation increased 10–20% when movements were made actively rather than passively, suggesting that efference copy signals (Bridgeman 1995) can also contribute to self-motion compensation. Another recent study reports compensation of approximately 75% during actively generated self-motion in the absence of visual self-motion cues (Dupin and Wexler 2013); passive self-motion was not investigated. Overall, such improvements in self-motion compensation with active self-motion suggest an important role of efference copy signals in object motion perception. Related to this, efference copy signals have been shown to improve the accuracy of reaching movements (Branch Coslett et al. 2008) and to facilitate accurate saccadic and pursuit eye movements (Lewis et al. 2001) after substantial proprioceptive loss.

Judging Object Motion with Visual Self-Motion Cues Only

A number of related studies have investigated the perception of object motion during visually simulated self-motion. A series of studies by Warren and Rushton (2007, 2008, 2009a, 2009b) has demonstrated that judgments of the trajectory of an object in 2-dimensional (2D) screen coordinates are biased when optic flow consistent with forward self-motion is presented. Remarkably, this occurs even when the optic flow is presented to one visual hemi-field, while the object is presented in the opposite hemi-field. These findings suggest that large-field visual mechanisms are involved in compensating for self-motion during judgments of object motion. The bias observed in these studies demonstrates that subjects did not judge object motion accurately in screen coordinates, but does not address the question of accuracy in world coordinates, because the stimuli and task were not sufficiently specified in 3D. Object distance was not represented, nor was the absolute speed of self-motion because the scene was devoid of scale information. Thus, the degree of compensation required for accurate perception of object motion in world coordinates was not specified.

We are aware of only one study that examined the accuracy of object trajectory estimates in world coordinates during visually simulated self-motion. Matsumiya and Ando (2009) presented subjects with large-field stereoscopic visual stimuli simulating forward self-motion through a room with textured walls. Subjects judged the 3D trajectory of a ball approaching their path from one side. Without simulated self-motion (i.e., textured room with no optic flow), trajectory judgments were made in retinal/screen coordinates. When self-motion was simulated visually, trajectory judgments were consistent with approximately 60% compensation for self-motion. This is of the same order as the 47% observed in our visual condition. Another recent study examining the accuracy of object speed judgments reports a similar value of 40% compensation during visual self-motion (Dupin and Wexler 2013).

Judging Object Motion During Combined Visual–Vestibular Stimulation

We found that world-referenced object motion judgments were most accurate when both visual and vestibular self-motion cues were present. Self-motion compensation in the combined condition (58%) was significantly greater than that seen in the visual (47%) and vestibular (0%) conditions. Dyde and Harris (2008) also demonstrated increased compensation for self-motion in the combined condition (∼90%), but only relative to a nonvisual condition (∼40%); they did not include a visual-only condition. More recently, Dupin and Wexler (2013) also compared the gain of self-motion compensation in visual, nonvisual (vestibular–proprioception–efference copy), and combined conditions during judgment of speed of object rotation. In agreement with our results and those of previous studies (Dyde and Harris 2008), compensation increased in the combined condition. Averaged across subjects, the combined compensation was roughly equal to the sum of the compensation observed in visual-only and nonvisual-only conditions (∼40% and ∼75% compensation, respectively), even though visual and nonvisual self-motion cues were placed in conflict.

Here, we consider 2 possible explanations for differences across self-motion conditions in our experiment that stem from distinct hypotheses about the origin of partial compensation. First, partial compensation may result from mis-estimation of self-motion velocity or object distance, as described above. In this case, increased compensation in the combined condition would result from more accurate estimates of self-motion speed and/or object distance. Prior research from our lab has shown that visual and vestibular cues interact during estimation of both self-motion (Fetsch et al. 2009) and distance (Dokka et al. 2011). However, these previous results do not directly quantify how visual–vestibular integration alters estimates of self-motion speed or object distance.

Secondly, partial compensation may reflect a mixed reliance on both observer-relative and world-relative estimates of object trajectory. In this case, degree of compensation may reflect the degree of belief in the underlying cause of the sensory signals, that is, causal inference (Kording et al. 2007). For example, if the observer believes that retinal image motion is caused by object motion alone, the observer-relative estimate may be used to achieve maximal sensitivity (as self-motion signals may carry noise), and there should be 0% compensation. On the other hand, if the observer believes that the pattern of retinal image motion is caused by a combination of object motion and self-motion, the world-relative estimate of object trajectory should be used, and 100% compensation should be observed. Partial compensation may be observed if there is uncertainty about the underlying cause of retinal image motion. If the observer believes that there is a 70% probability that the retinal image motion is caused by combined object motion and self-motion, the observer could select the world-centered estimate 70% of the time and the observer-centered estimate 30% of the time, a scheme that we refer to as Model Sampling. Alternatively, an estimate of object trajectory may be computed as a weighted average of observer-relative and world-relative estimates with weights depending on the relative probabilities of the 2 causes (0.3 and 0.7, respectively, in this example); we refer to this as Model Averaging. Note that Model Averaging requires observer-relative and world-relative estimates to be in the same units, just expressed in different reference frames. We assume that the

egocentric estimate is defined as angular deviation of the object trajectory from vertical in the head-fixed frame, and the allocentric estimate as angular deviation from vertical in the world-fixed frame. A final possibility, Model Selection, dictates that the observer simply selects the alternative for which the probability is >50%.

In principle, either Model Averaging or Model Sampling could account for the partial compensation observed in our data. To explore these possibilities, we performed simulations as described in detail in the Supplementary Materials. These simulations reveal that Model Averaging is consistent with a monotonically increasing relationship between percent compensation and psychophysical thresholds (Supplementary Fig. 5), similar to what we observe in our data. In contrast, Model Sampling predicts psychophysical thresholds that are maximal for intermediate probabilities of a world-centered cause. Although the uncertainty in these predictions does not allow us to firmly rule out Model Sampling, our data appear most consistent with the predictions of a Model Averaging scheme. Our data cannot be explained by Model Selection, as this scheme does not predict partial compensation.

If we assume that our data reflect causal inference-based Model Averaging, then percent compensation provides a metric of the observers' estimates of probabilities associated with the alternative interpretations of the stimuli. In the vestibular condition, we observed close to 0% compensation, suggesting that the observer interprets this constellation of stimulus features (sparse, observer-fixed visual scene) to be most likely associated with object motion only, despite passive vestibular stimulation. This suggests that the observer-fixed visual frame provided by the screen boundaries and fixation target have a powerful influence on observers' interpretation of the scene, in the absence of decision-related feedback. In the visual condition, the full-field optic flow pattern makes the self-motion interpretation more likely, leading to increased compensation (47%). And in the combined condition, congruent optic flow and vestibular stimulation make the self-motion interpretation even more likely (58% compensation).

The neural mechanisms by which integration of visual and vestibular self-motion signals improves the accuracy of world-relative object motion judgments are presently unclear. However, we have previously speculated that a subgroup of multisensory neurons in the area MSTd with opposite visual and vestibular heading tuning might contribute to dissociating object motion from self-motion (Gu et al. 2006, 2008; Morgan et al. 2008). Indeed, preliminary results show that addition of vestibular self-motion information stabilizes the object motion tuning of opposite cells when self-motion direction varies, which is not the case for MSTd cells with congruent visual and vestibular heading tuning (Sasaki et al. 2012). Therefore, opposite cells might provide a neural substrate for the effects we report here, although it is currently unknown how feedback training may modulate these neuronal responses. Further studies in macaques trained to judge world-relative and observer-relative object motion will be critical to address these open issues.

Other Factors Influencing Degree of Compensation

We found that percent compensation is influenced not only by the stimulus modalities involved, but also by other stimulus parameters such as the speed of self-motion (Fig. 4E,F). Specifically, compensation decreased with increasing speed of self-motion. We assume that faster self-motion would increase the probability that the retinal image results from a combination of object motion and self-motion. Thus, in the causal inference hypothesis outlined above, increased self-motion speed should be associated with greater percentage compensation, not less. Alternatively, the decrease in compensation with increasing self-motion speed may be attributed to subjects' under-estimating self-motion speed, with the degree of underestimation increasing with the true speed. While there was similar improvement in compensation at all self-motion velocities after feedback training, compensation was still poorer at faster self-motion speeds, indicating that mis-estimation remained correlated with the true speed even after training.

Factors such as speed of self-motion may explain some differences across studies in the literature. For example, Dyde and Harris (2008) observed approximately 90% compensation in their combined condition compared with 58% in the present study. However, the peak velocity of self-motion in their study was approximately 10 cm/s, lower than the minimum peak velocity used in our study. At 12 cm/s, we observed >80% compensation in the combined condition (Fig. 4E, leftmost red bar), thus making our results compatible with those of Dyde and Harris (2008).

We also investigated the possible influence of eye movements on self-motion compensation. In our experiments, observers were instructed to fixate a head-fixed fixation point. There was no correlation between mean eye velocity and bias in perceived object trajectory. However, as eye movements can play a key role in self-motion perception (Royden et al. 1992; Banks et al. 1996; Crowell et al. 1998), it is possible that partial suppression of the vestibulo-ocular reflex during fixation may have influenced perceptual judgments. However, in previous studies, observers were instructed to track the moving object with their eyes (Dyde and Harris 2008; Matsumiya and Ando 2009) and exhibited perceptual behavior similar to our results, suggesting that eye movements are unlikely to account for observers' behavior when judging object motion.

We found a modest, but significant, effect of self-motion direction on percent compensation, such that less compensation was observed during leftward than rightward self-motion. Although our analysis shows no significant effect of the starting position of the object on object trajectory judgments, the range of ending positions of the object does differ for the 2 directions of self-motion. This difference in ending position could potentially account for the direction-dependent component of bias. Note, however, that our main findings hold for both directions of self-motion, so this effect does not compromise our conclusions.

Accuracy Versus Precision of Object Trajectory Judgments

We observed a positive correlation between percent compensation and threshold, suggesting that increased accuracy comes at the cost of decreased precision. Compensation was least, and thresholds lowest, in the vestibular condition, while compensation was greatest and thresholds highest in the combined condition. Also, thresholds increased with self-motion speed, most likely due to an increase in signal-dependent noise (i.e., Weber's law). On the surface, the present finding of higher thresholds in the combined condition than in the visual condition appears to conflict with recently published findings from our laboratory

([MacNeilage et al. 2012](#)). However, in that experiment, observers were not asked to make judgments in world coordinates and accuracy was not assessed. If degree of compensation were not changing across stimulus conditions, it is reasonable to expect better performance in the combined condition when self-motion information is integrated across sensory modalities.

The correlation between percent compensation and threshold was also observed when evaluating the effect of feedback training on object trajectory judgments. We demonstrated that feedback training in the vestibular condition allows subjects to compensate near-perfectly for self-motion, whereas there was essentially no compensation in this condition before feedback training. This effect generalized to untrained velocities and stimulus modalities, suggesting strongly that feedback allowed subjects to learn to judge object motion accurately in world coordinates. In the causal inference scheme, this suggests that probabilities associated with alternative causal interpretations may be learned through feedback. At the same time, the increased accuracy gained through feedback training was associated with decreased precision (elevated thresholds) in all stimulus conditions. This effect echoes the correlation observed before feedback training.

What accounts for the correlation between compensation and precision? There is less noise associated with object motion estimates made in observer coordinates. Observer-centered judgments simply require evaluating visual motion relative to the observer-fixed visual features, such as the fixation point and display frame. On the other hand, more noise is associated with judgments in world coordinates because judgments in this reference frame involve subtracting the estimated self-motion from the retinal image motion of the object. Thus, variability in judgments of object motion in the world must incorporate variability from both processing of self-motion signals (visual and/or vestibular) and processing of retinal image motion.

Degraded precision when combining multimodal signals for perceptual judgments in world coordinates has been previously reported in studies of reaching movements ([Sober and Sabes 2003](#), [2005](#); [Burns and Blohm 2010](#)), visual object localization ([Burns et al. 2011](#)), and spatial working memory ([Golomb and Kanwisher 2012a](#), [2012b](#)). [Burns and Blohm (2010)](#) found that head roll away from the upright position increased variability of arm reaches. [Burns et al. (2011)](#) found similar influences of head roll-dependent noise on visual object localization. They observed that eccentric head roll increases the variability in subjects' ability to discriminate visual object location, as indicated by increased thresholds. [Golomb and Kanwisher (2012a)](#) tested a spatial working memory task in which subjects remembered a cued location in spatiotopic or retinotopic coordinates while making saccadic eye movements during the memory delay. Subjects were more precise at reporting retinotopic rather than spatiotopic location. Our results demonstrating increased thresholds when object trajectory judgments were made in world coordinates are consistent with these previous findings. In our experiment, 2 sources of noise may have contributed to the elevated thresholds associated with world-centered judgments. The first source is the noise inherent in visual and vestibular self-motion estimates. Incorporation of these noisy signals is required for world-centered judgments and could have increased the noise in the final object trajectory estimates. The second source of noise could be the reference frame transformation required for combining multimodal signals. Vestibular signals are encoded in head coordinates, whereas visual motion

signals are encoded in retinal coordinates. To judge object trajectory in world coordinates, both sensory signals would have to be transformed into a common allocentric reference frame and such a transformation of signals may contribute additional noise ([Sober and Sabes 2003](#), [2005](#); [Burns and Blohm 2010](#); [Burns et al. 2011](#)).

Given that there is a cost (in terms of precision) to representing object motion in world coordinates, it may be beneficial for the brain to retain flexibility in how object motion is represented. When a self-centered reference frame is appropriate (see Introduction), ignoring self-motion and judging object motion in observer coordinates can maximize precision. In such situations, it may be irrelevant that the estimate of object motion in world coordinates is inaccurate. In other situations, accurate estimation of object motion in the world may be critical. These considerations may help explain the default tendency of observers to make judgments in observer coordinates in our vestibular condition. As expected from causal inference theory, when there is uncertainty about the most appropriate reference frame, the brain may resort to using a weighted combination of egocentric and allocentric estimates with weights that depend on the amount of evidence favoring one reference frame over the other. This can potentially explain the precision–accuracy tradeoff observed here: the higher the weight given to the world-centric estimate, the more noise will be added to the final object motion estimate.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxford-journals.org/.

Funding

Notes

Conflict of Interest: None declared.

References

Banks MS, Ehrlich SM, Backus BT, Crowell JA. 1996. Estimating heading during real and simulated eye movements. Vision Res. 36:431–443.

Blohm G, Missal M, Lefevre P. 2005. Processing of retinal and extraretinal signals for memory-guided saccades during smooth pursuit. J Neurophysiol. 93:1510–1522.

Blohm G, Missal M, Lefevre P. 2003. Smooth anticipatory eye movements alter the memorized position of flashed targets. J Vis. 3:761–770.

Blohm G, Optican LM, Lefevre P. 2006. A model that integrates eye velocity commands to keep track of smooth eye displacements. J Comput Neurosci. 21:51–70.

Branch Coslett H, Buxbaum LJ, Schwoebel J. 2008. Accurate reaching after active but not passive movements of the hand: evidence for forward modeling. Behav Neurol. 19:117–125.

Brenner E, Smeets JB, van den Berg AV. 2001. Smooth eye movements and spatial localisation. Vision Res. 41:2253–2259.

Bridgeman B. 1995. A review of the role of efference copy in sensory and oculomotor control systems. Ann Biomed Eng. 23:409–422.

Burns JK, Blohm G. 2010. Multi-sensory weights depend on contextual noise in reference frame transformations. Front Hum Neurosci. 4:221.

Burns JK, Nashed JY, Blohm G. 2011. Head roll influences perceived hand position. J Vis. 11(9):3, 1–9.

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.

Crowell JA, Banks MS, Shenoy KV, Andersen RA. 1998. Visual self-motion perception during head turns. Nat Neurosci. 1:732–737.

Daye PM, Blohm G, Lefevre P. 2010. Saccadic compensation for smooth eye and head movements during head-unrestrained two-dimensional tracking. J Neurophysiol. 103:543–556.

Dessing JC, Crawford JD, Medendorp WP. 2011. Spatial updating across saccades during manual interception. J Vis. 11(10):4, 1–18.

Dokka K, MacNeilage PR, DeAngelis GC, Angelaki DE. 2011. Estimating distance during self-motion: a role for visual-vestibular interactions. J Vis. 11(13):2, 1–16.

Dupin L, Wexler M. 2013. Motion perception by a moving observer in a three-dimensional environment. J Vis. 13(2):15, 1–14.

Dyde RT, Harris LR. 2008. The influence of retinal and extra-retinal motion cues on perceived object motion during self-motion. J Vis. 8 (5):1–10.

Fajen BR, Matthis JS. 2013. Visual and non-visual contributions to the perception of object motion during self-motion. PLoS One. 8: e55446.

Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE. 2011. Neural correlates of reliability-based cue weighting during multisensory integration. Nat Neurosci. 15:146–154.

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.

Gogel WC. 1981. The role of suggested size in distance responses. Percept Psychophys. 30:149–155.

Gogel WC. 1982. Analysis of the perception of motion concomitant with a lateral motion of the head. Percept Psychophys. 32:241–250.

Gogel WC. 1990. A theory of phenomenal geometry and its applications. Percept Psychophys. 48:105–123.

Gogel WC, Tietz JD. 1992. Determinants of the perception of sagittal motion. Percept Psychophys. 52:75–96.

Gogel WC, Tietz JD. 1977. Eye fixation and attention as modifiers of perceived distance. Percept Mot Skills. 45:343–362.

Golomb JD, Kanwisher N. 2012a. Retinotopic memory is more precise than spatiotopic memory. Proc Natl Acad Sci USA. 109: 1796–1801.

Golomb JD, Kanwisher N. 2012b. Higher level visual cortex represents retinotopic, not spatiotopic, object location. Cereb Cortex. 22: 2794–2810.

Gray R, Macuga K, Regan D. 2004. Long range interactions between object-motion and self-motion in the perception of movement in depth. Vision Res. 44:179–195.

Gu Y, Angelaki DE, DeAngelis GC. 2008. Neural correlates of multisensory cue integration in macaque MSTd. Nat Neurosci. 11: 1201–1210.

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.

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.

Kording KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L. 2007. Causal inference in multisensory perception. PLoS One. 2: e943.

Lewis RF, Zee DS, Hayman MR, Tamargo RJ. 2001. Oculomotor function in the rhesus monkey after deafferentation of the extraocular muscles. Exp Brain Res. 141:349–358.

MacNeilage PR, Zhang Z, Angelaki DE. 2012. Vestibular facilitation of optic flow parsing. PLoS One. 7:e40264.

Matsumiya K, Ando H. 2009. World-centered perception of 3d object motion during visually guided self-motion. J Vis. 9(15):11–13.

Mesland BS, Finlay AL, Wertheim AH, Barnes GR, Morland AB, Bronstein AM, Gresty MA. 1996. Object motion perception during egomotion: patients with a complete loss of vestibular function vs. normals. Brain Res Bull. 40:459–465.

Morgan ML, DeAngelis GC, Angelaki DE. 2008. Multisensory integration in macaque visual cortex depends on cue reliability. Neuron. 59:662–673.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1992. Numerical Recipes in C. 2nd ed. New York: Cambridge University Press. pp. 650–655.

Royden CS, Banks MS, Crowell JA. 1992. The perception of heading during eye movements. Nature. 360:583–585.

Sasaki R, Angelaki DE, DeAngelis GC. 2012. Role of vestibular signals in dissociating self-motion and object motion in macaque area MSTd. Program No. 467.05. 2012 Neuroscience Meeting Planner. New Orleans, LA: Society for Neuroscience, Online.

Savitzky A, Golay MJE. 1964. Smoothing and differentiation of data by simplied least squares procedures. Anal Chem. 36:1627–1639.

Shebilske WL, Proffitt DR. 1981. The priority of perceived distance for perceiving motion has not been demonstrated: critical comments on gogel's "the sensing of retinal motion". Percept Psychophys. 29:170–172.

Sober SJ, Sabes PN. 2003. Multisensory integration during motor planning. J Neurosci. 23:6982–6992.

Sober SJ, Sabes PN. 2005. Flexible strategies for sensory integration during motor planning. Nat Neurosci. 8:490–497.

Swanston MT, Wade NJ. 1988. The perception of visual motion during movements of the eyes and of the head. Percept Psychophys. 43:559–566.

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.

Van Pelt S, Medendorp WP. 2008. Updating target distance across eye movements in depth. J Neurophysiol. 99:2281–2290.

Wallach H. 1987. Perceiving a stable environment when one moves. Annu Rev Psychol. 38:1–27.

Warren PA, Rushton SK. 2007. Perception of object trajectory: parsing retinal motion into self and object movement components. J Vis. 7 (2):1–11.

Warren PA, Rushton SK. 2008. Evidence for flow-parsing in radial flow displays. Vision Res. 48:655–663.

Warren PA, Rushton SK. 2009a. Optic flow processing for the assessment of object movement during ego movement. Curr Biol. 19:1555–1560.

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.

Wexler M. 2003. Voluntary head movement and allocentric perception of space. Psychol Sci. 14:340–346.

Wichmann FA, Hill NJ. 2001a. The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys. 63:1293–1313.

Wichmann FA, Hill NJ. 2001b. The psychometric function: II. Bootstrap-based confidence intervals and sampling. Percept Psychophys. 63:1314–1329.

Wyatt HJ. 1998. Detecting saccades with jerk. Vision Res. 38:2147–2153.