

Experience affects the use of ego-motion signals during 3D shape perception

Anshul Jain

SUNY Eye Institute and Graduate Center for Vision Research,
SUNY College of Optometry, New York, NY, USA



Benjamin T. Backus

SUNY Eye Institute and Graduate Center for Vision Research,
SUNY College of Optometry, New York, NY, USA



Experience has long-term effects on perceptual appearance (Q. Haijiang, J. A. Saunders, R. W. Stone, & B. T. Backus, 2006). We asked whether experience affects the appearance of structure-from-motion stimuli when the optic flow is caused by observer ego-motion. Optic flow is an ambiguous depth cue: a rotating object and its oppositely rotating, depth-inverted dual generate similar flow. However, the visual system exploits ego-motion signals to prefer the percept of an object that is stationary over one that rotates (M. Wexler, F. Panerai, I. Lamouret, & J. Droulez, 2001). We replicated this finding and asked whether this preference for stationarity, the “stationarity prior,” is modulated by experience. During training, two groups of observers were exposed to objects with identical flow, but that were either stationary or moving as determined by other cues. The training caused identical test stimuli to be seen preferentially as stationary or moving by the two groups, respectively. We then asked whether different priors can exist independently at different locations in the visual field. Observers were trained to see objects either as stationary or as moving at two different locations. Observers’ stationarity bias at the two respective locations was modulated in the directions consistent with training. Thus, the utilization of extraretinal ego-motion signals for disambiguating optic flow signals can be updated as the result of experience, consistent with the updating of a Bayesian prior for stationarity.

Keywords: perceptual learning, structure from motion, ego-motion, extraretinal signals, perceptual bias, cue recruitment

Citation: Jain, A., & Backus, B. T. (2010). Experience affects the use of ego-motion signals during 3D shape perception. *Journal of Vision*, 10(14):30, 1–14, <http://www.journalofvision.org/content/10/14/30>, doi:10.1167/10.14.30.

Introduction

Knowledge of surroundings is essential to most organisms. This knowledge is acquired both by passive observation of the environment and active interaction with it. Humans often use active and passive acquisitions in conjunction, sometimes by integrating information from different sensory modalities. For example, the sharpness of a pencil tip is judged both by sight and tactile exploration. In this case, multiple senses provide information about the same object property that may be combined in a statistically optimal fashion (e.g. Adams, Graf, & Ernst, 2004; Alais & Burr, 2004; Ernst & Banks, 2002). In other cases, signals from one modality are used to interpret or disambiguate signals from another modality rather than being averaged with them. For example, auditory signals can disambiguate whether two disks are seen as bouncing off or gliding past each other (Sekuler, Sekuler, & Lau, 1997; Watanabe & Shimojo, 2001). In the current study, we examine how ego-motion signals are used to disambiguate visually acquired information about 3D shape, with the following question in mind: Can learning change how ego-motion signals are used during visual perception?

Optic flow is the motion pattern generated within the optic array (and thus on the retina) by relative motion between the observer and objects in the world. Optic flow is a potent cue to depth or 3D shape of an object (Domini & Caudek, 2003; Koenderink, 1986; Koenderink & van Doorn, 1975; Longuet-Higgins & Prazdny, 1980; Sperling, Doshier, & Landy, 1990; Wallach & O’Connell, 1953). Motion parallax is a special case of optic flow, in which the speed of each point depends on its depth in the scene. Thus, parallax carries information about the 3D structure of the object. Perceptual recovery of this structure is called structure from motion (Rogers & Graham, 1979). The motion parallax can be generated either by observer motion relative to a stationary object in space or by object motion relative to a stationary observer. The optic flow in these two cases is the same. Thresholds in structure-from-motion displays depend only on the relative motion between observer and object (Rogers & Graham, 1979, 1982), and it has often been assumed that structure from motion operates similarly for observer and object motion.

However, it was recently shown that identical optic flow patterns are interpreted differently depending on whether they were generated by observer or object motion (Fantoni, Caudek, & Domini, 2010; van Boxtel, Wexler, & Droulez, 2003; Wexler, Panerai, Lamouret, & Droulez,

2001). An observer who uses image flow from a device that fixates the object, such as an eye might do, will find that the image flow is ambiguous. Absent additional depth cues (such as stereo or relative brightness), there are two objects that are consistent with the motion parallax. One is the object itself and the other is the object's depth-reversed mirror image, deforming to accord with perspective cues, where the parallax was caused by the opposite translational motion between the observer and the object. However, when motion parallax is generated by observer motion, the visual system prefers the percept that is consistent with a stationary object in the allocentric frame of reference. This disambiguation can be achieved using extraretinal signals generated during ego-motion (Wexler, Panerai et al., 2001) together with an assumption, presumably based on experience, that objects are likely to be stationary in the world (neither translating nor rotating relative to the local environment). The ego-motion signals specify the translation of the observer relative to the (non-translating) object. After this motion is known, one of the two structure-from-motion interpretations is an object that does not rotate relative to the world, while the other is a depth-reversed object with an angular velocity relative to the world that is twice the rate of change in the angle of regard to the object.¹ At this point, the perceptual system can exert its preference for the non-rotating object.

The human visual system presumably draws on experience to interpret complex and/or ambiguous signals. Sensory signals contain measurement noise, and even noise-free signals could be consistent with multiple interpretations of the world (Helmholtz, 1910/1925). The perceptual system can rely on prior experience to choose between the possible perceptual constructs (Haijiang, Saunders, Stone, & Backus, 2006; Sinha & Poggio, 1996). The influence of pre-existing (prior) belief on perception, measured as a perceptual bias, has often been studied in a Bayesian framework (e.g., Kersten, Mamassian, & Yuille, 2004; Langer & Bulthoff, 2001; van Ee, Adams, & Mamassian, 2003; Weiss, Simoncelli, & Adelson, 2002), and it is reasonable to suppose that experience contributes to this belief. In this framework, priors based on previous experience are combined with the cues present in the current sensory input to arrive at a percept. Knowledge of environmental statistics is important for interpreting sensory input; for example, implicit knowledge that light usually comes from above is used during the perceptual recovery of 3D object shapes from shading cues (Sun & Perona, 1998).

The light-from-above prior can be modified by experience (Adams et al., 2004), as can certain prior “beliefs” that the perceptual system has about shape (Champion & Adams, 2007; Sinha & Poggio, 1996). Prior belief about the appropriate weighting for different visual cues can also be modified, after experience of different cue reliabilities during experimental training (Ernst, Banks, & Bulthoff, 2000; Jacobs & Fine, 1999; Knill, 2007). Our question therefore is whether the perceptual system's prior belief

about an object's stationarity, the “stationarity prior” (Wexler, Lamouret, & Droulez, 2001), is fixed or can be modulated based on experience. This question was raised by Wallach, Stanton, and Becker (1974), who believed that it was impossible to demonstrate such a modulation and it has remained unanswered since. The stationarity prior is an interesting case, insofar as it determines the utilization of ego-motion signals for interpreting motion parallax signals. Modulation of this prior would affect a computation in which two independent measures of different physical properties (ego-motion, from extraretinal signals, and optic flow, from retinal signals) are combined across sensory modalities to estimate the 3D structure of an object.

A second objective of the current study was to determine whether more than one stationarity prior can be learned for different locations in the visual field. Different stable biases for other perceptual ambiguities can simultaneously exist at different locations in the visual field (Backus & Haijiang, 2007; Carter & Cavanagh, 2007; Harrison & Backus, 2010a). To the extent that different neural machinery resolves the ambiguity at different locations within the visual field, separate stationarity biases might in fact be trainable at different locations. If different priors can be learned at different locations, it would mean that some or all of the learning occurs at a level within the visual system that has access to spatial location. Our design does not distinguish learning of this location-specific bias in a retinotopic frame of reference from learning in an allocentric frame of reference.

We conducted three experiments to examine the effect of training on the stationarity prior. The first experiment replicated the study of Wexler, Panerai et al. (2001). The second experiment tested whether the stationarity prior is updated in response to a change in environmental statistics: we manipulated the frequency of object stationarity during training and measured whether this caused differences in perceptual bias. In the third experiment, we examined whether the prior is expressed as a single uniform bias across the whole visual field or if different priors can be learned at different locations in the visual field respectively.

General methods

Observers

Forty-two observers participated in the study: six in Experiment 1, seventeen in Experiment 2, and nineteen in Experiment 3. The data from five observers, three from Experiment 3 and one each from Experiments 1 and 2, were discarded because they could not perform the task. All observers were naive to the purposes of the experiments. The experiments were conducted in compliance with the standards set by the IRB at SUNY State College of Optometry. Observers gave their informed consent

prior to their inclusion in the study and were paid for their participation. All observers had normal or corrected-to-normal vision.

Apparatus

The experiments were implemented on a Dell Precision 3400 computer (Windows platform) using the Python-based virtual reality toolkit Vizard 3.0 (WorldViz LLC, Santa Barbara, CA, USA). Visual stimuli were back-projected on a screen ($2.4\text{ m} \times 1.8\text{ m}$) using a Christie Mirage S+ 4K projector. The display refresh rate was fixed at 120 Hz and the screen resolution was 1024×768 pixels. Head position was tracked using Hiball-3100 (3rdTech Durham, NC, USA) wide-area precision tracker. The head tracker monitored observers' head positions and provided the head coordinates at a minimum rate of 1400 Hz to the computer used to generate the stimuli. Sensor location (x, y, z) was correct to within 1 mm, with similar trial-to-trial precision; orientation (3 DOF) was correct to within 1 deg. The tracker itself had a latency of less than 1 ms; the computer and display were updated 30–35 ms after a head movement, as measured using a high-speed (1000 Hz) video camera. The stimuli on the screen were updated based on the head position of the observer. Observers were seated at a distance of 2 m from the projection screen and wore an eye patch so as to view stimuli monocularly using their dominant eye.

Stimuli

Stimuli consisted of a planar grid with the slant fixed at 45° , similar to the one used by Wexler, Panerai et al.

(2001). The tilt of the grid was varied randomly between trials. Slant is defined as the angle between the surface normal and the horizontal line of sight to that surface. It is the angle between the surface and the frontoparallel plane. Tilt is given by the angle of the projected surface normal in the frontoparallel plane. Observer's task was to report the tilt of the grid using a gauge figure. The gauge figure was made up of a disk with a pointer normal to the surface of the disk. The disk was 10 cm in radius and the pointer needle was 20 cm in length for Experiments 1 and 2. The corresponding dimensions were halved to 5 cm and 10 cm, respectively, for Experiment 3. The disk had a slant of 45° and observers could turn the disk and orient the pointer needle to match the perceived tilt of the grid. The grid was derived from a 10×10 array of squares. Each square had a side length of 5 cm in Experiments 1 and 2. The side length was halved to 2.5 cm for Experiment 3. Thus, the grid subtended a visual angle of 20° in the first two experiments and a visual angle of 10° in Experiment 3.

The grid stimuli were presented in two configurations, no-conflict stimulus configuration and conflict stimulus configuration (Figure 1). In Experiment 1, the object was always simulated as stationary, although sometimes it was perceived to move. In Experiments 2 and 3, the object was simulated as stationary on some trials and rotating on other trials. The no-conflict stimulus was constructed as described above. It had a 45° slant and a random tilt. In the conflict configuration, the stimulus was constructed by performing a perspective projection (approximately orthographic in Experiment 3, explained later) of the no-conflict stimulus onto a plane passing through the center of the grid. That plane had the same slant, but the tilt differed from the no-conflict tilt by a fixed angle (45, 90, 135, or 180 deg in Experiment 1; 180 deg in Experiments 2 and 3). Observers viewing this stimulus could in principle perceive

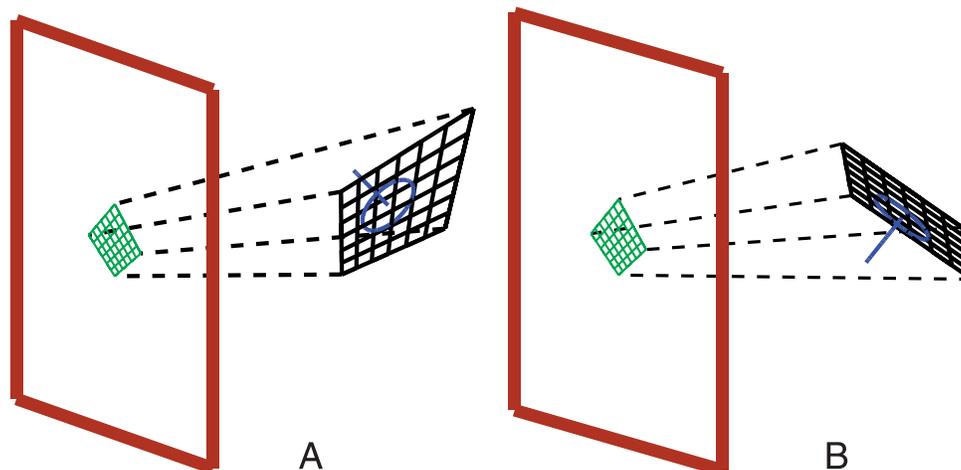


Figure 1. Stimuli used in the experiments. (A) Conflict stimulus: Virtual object is irregular with an upward tilt. The perspective cues signal a downward tilt while the optic flow cues (under ego-motion with stationarity) signal an upward tilt. (B) No-conflict stimulus: Perspective and optic flow both signal a downward tilt.

either of the two (opposite) tilts that were consistent with motion parallax, or the tilt specified by perspective, or something intermediate. If they used parallax to recover tilt, they would be expected to perceive a rigid object that was irregular in construct (a distorted grid). If they used perspective, they would perceive a more regular grid object that was non-rigid due to motion cues (see Wexler, Panerai et al., 2001). In the special case of 180-deg conflict, perspective agrees with the dual of the actual object so an apparently no-conflict object could be perceived; however, if the actual object were simulated to be stationary, this dual would be seen to rotate, and vice versa (as occurs in the hollow mask illusion, e.g., Papatomas & Bono, 2004).

On any given trial in [Experiment 1](#), only ten squares from the grid (randomly selected) were made visible to the observer, for consistency with the study by Wexler, Panerai et al. (2001). We did this to reduce the strength of perspective cues by avoiding presenting long converging lines while still giving an impression of a surface. Further, with only a few squares visible the irregularity of the grid was probably less noticeable when observers extracted the shape using motion cues (in conflict with perspective cues). In [Experiments 2](#) and [3](#), an additional eight squares were also made visible. The eight squares were chosen such that the four corner squares and one of two squares adjacent to each corner square were also made visible. This was done to ensure that the stimuli subtended the same visual angle on every trial irrespective of the squares chosen.

Experiment 1—Structure from motion and ego-motion

[Experiment 1](#) replicated the apparent tilt experiment of Wexler, Panerai et al. (2001). In this experiment, perspective cues that specified surface tilt were put into conflict with motion parallax cues, in order to examine the effect of ego-motion signals on the use of parallax. The conflict between the tilts specified by parallax and perspective was randomly chosen from 45, 90, 135, or 180 degrees. On any given trial, the tilt for one interpretation of the parallax signal differed from the dual interpretation by 180 deg. Thus, a conflict of 45 deg between perspective and one parallax interpretation was also a conflict of 225 deg (or -135 deg) between perspective and the dual parallax interpretation.

The apparent tilt of the depicted surface revealed how much weight the observer gave to parallax, and whether they preferred the stationary interpretation. If apparent tilt agrees with the stationary interpretation, then ego-motion signals are used to construct object appearance and we then predict two things. First, since observers prefer the stationary interpretation of the parallax cue, the biasing effect of parallax (away from slant as specified by

perspective) should be in the direction of the stationary surface's slant, rather than that of the dual (180 deg away). Second, the observer's percept should be based more on parallax, and less on perspective, when the parallax is caused by ego-motion than if the same visual stimulus is presented to the passive observer.

Under the current experimental procedure, when observers estimated the stimulus tilt based on parallax, they presumably perceived a rigid but non-regular object (with veridical tilt or inverted tilt), but when they computed the stimulus tilt based on perspective, they perceived a more regular object that deformed non-rigidly. This was true both when the optic flow was generated by observer motion (*active* trials) and object motion (*passive* trials). Critically, however, on active trials the simulated object when perceived based on parallax cues was also stationary in an allocentric frame of reference when perceived with veridical tilt, unlike when it was perceived based on perspective cues or based on parallax cues but with inverted tilt. Therefore, the rigidity assumption would predict more parallax-based responses than perspective-based responses on both active and passive trials. In summary, if the human perceptual system relies on the stationarity assumption over and above the rigidity assumption we should observe a stronger preference for parallax during active trials as compared to passive trials together with fewer inverted responses on active trials.

Experimental procedure

The experimental procedure was similar to that of Wexler, Panerai et al. (2001). The experiment consisted of two kinds of trials, active trials and passive trials. On active trials, observers swayed from side to side as they viewed the stimulus to generate motion parallax. The stimulus presented on the screen was a perspective projection of the slanted planar grid (defined above) computed based on the observer's current eye location (the eye's orientation at that location being irrelevant). The sensor was located at the top of the helmet, so eye position was computed by adding a fixed displacement vector (measured for one adult) to the sensor's current position. Eye locations in (x, y, z) coordinates were tracked during each active trial to record the observers' viewpoints and were then used to generate passive trials. During passive trials, observers were instructed to remain stationary but viewed the same optic flow as on one of the active trials. Optic flow on passive trials did not replicate the stimulus on the display screen but rather the flow as measured at the eye (i.e., it was dynamically recomputed to be correct for current eye location within the room). Hence, a difference between active and passive trials was the extraretinal signals about ego-motion; the optic array was kept constant (as was retinal input, assuming accurate fixation). Observers were instructed to fixate on a red dot

presented at the center of the grid during the entire trial and fixation was not otherwise monitored.

A beep sound was played at the beginning of each active trial, indicating to the observer to start their movement cycle. Observers were asked to move their head laterally starting from upright position to the left until they heard a beep sound and then to the right until they heard another beep sound and back to the center two times. The beep was played when the observers' head deviated from the upright position by 20 cm, thus controlling the amplitude of head movement. To prevent observers from seeing the grid while stationary, the stimulus appeared on the screen after the first half-cycle of movement (center to left and back to the center). Observers were instructed to move their head using their torso and to minimize head tilts, which minimized the curl component of the optic flow (Koenderink, 1986) that does not provide information about the 3D structure. They were also instructed to keep their motion as smooth and as continuous as possible. The grid was centered in front of the eye. It disappeared after the observer completed the two movement cycles (simulated movements on passive trials) and the gauge figure was then presented on the screen. Observers rotated the gauge figure (about the line of sight) by pressing either of two keys until the pointer needle was aligned with the remembered perceived tilt of the grid.

On any given trial, active or passive, the stimulus could be presented in either of the two configurations (conflict or no conflict). The conflict tilt (ΔT) was chosen pseudo-randomly and could take any one of five values (0° , 45° , 90° , 135° , 180°). A ΔT of 0° conflict corresponded to

no-conflict stimulus configuration. Active and passive trials were presented in alternate blocks of 50 trials each. Observers participated in eight blocks (four blocks each of active and passive trials), spread over two sessions. Each session took approximately 1 h.

Results

Figure 2 shows the histogram of reported tilts relative to the parallax-defined tilt for all observers. Hence, reported tilts with a value of 0° correspond to the responses based on the tilt defined by motion cues under stationary interpretation, while reported tilts with a value equal to ΔT (45° , 90° , 135° , 180°) correspond to the responses based on the tilt defined by perspective cues. With these large conflicts, it would not be surprising if the perceptual outcome on each trial were one of the two possible tilts rather than an average (Landy, Maloney, Johnston, & Young, 1995). The bimodal nature of the distributions is clearly evident in the histograms (and in individual data).

We classified the responses that were $\pm 22.5^\circ$ from 0° and 180° as “parallax” responses. The proportion of trials based on parallax includes both the trials when the simulated tilt was perceived as well as the trials when the opposite tilt was perceived. The responses that were within $\pm 22.5^\circ$ of ΔT were classified as “perspective” responses. For ΔT of 0° and 180° , it is not possible to classify responses as parallax and perspective responses, so we have excluded them from the data analysis. We conducted a three-way repeated measures ANOVA with

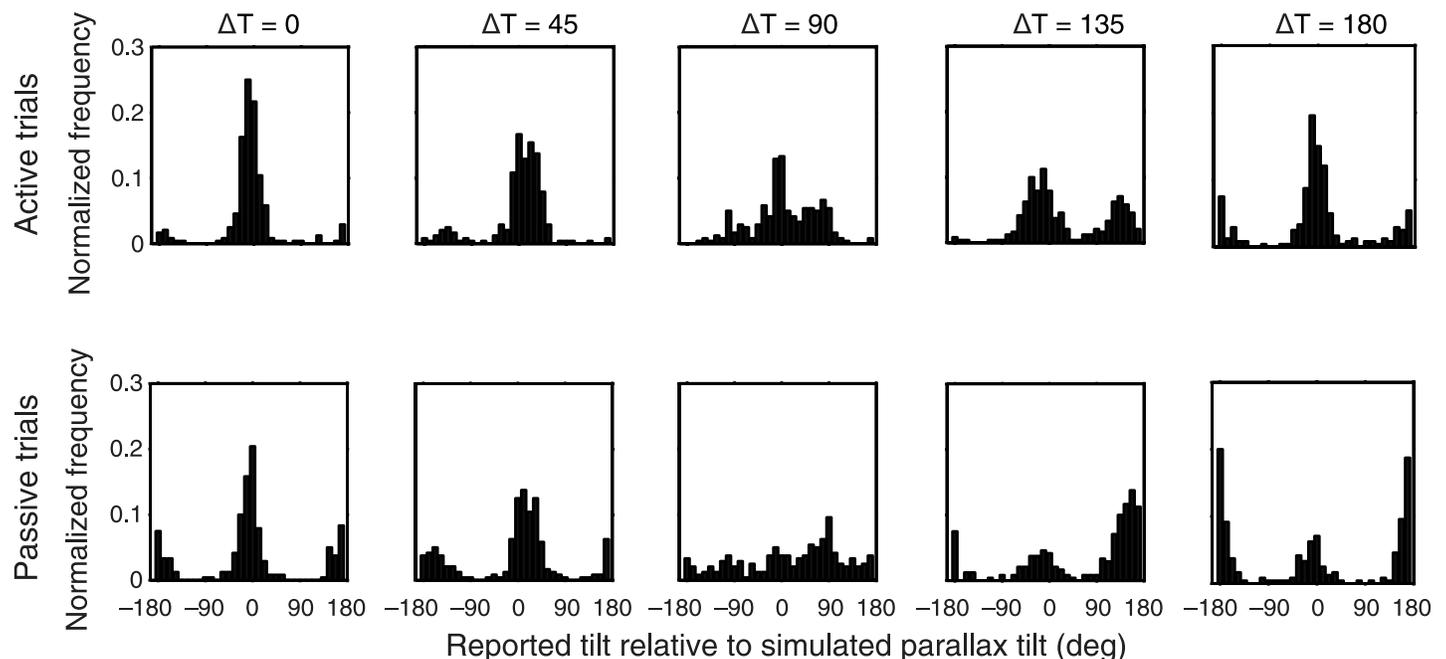


Figure 2. Histograms show the proportion of trials as a function of reported tilt for different values of ΔT . The top row corresponds to active trials and the bottom row corresponds to passive trials. Response tilt = 0° and 180° corresponds to motion cue-defined tilts while response tilt = ΔT corresponds to perspective cue-defined tilt. Data replicate of those of Wexler, Panerai et al. (2001).

proportion of trials as the dependent measure and trial type (active or passive), response type (parallax cue based or perspective cue based), and tilt conflict as independent variables (Wexler, Panerai et al., 2001). We observed a significant effect of the response type ($F(1,9) = 13.361$, $p \ll 0.001$): overall, observers' responses were based on parallax more often than perspective. There was also a significant effect of tilt conflict ($F(2,18) = 25.66$, $p \ll 0.001$); [Figure 2](#) shows that for tilt conflict = 45° , the peaks of the two distributions are merged, leading to higher proportions for both type of responses. Post-hoc analysis revealed that the proportions were significantly higher for conflict tilt = 45° , but there was no difference for the other two conflicts.

The most important result was a significant interaction ($F(2,18) = 9.49$, $p < 0.01$) between response type (parallax cue based or perspective cue based) and trial type (active or passive). Observers perceived the stimulus based on parallax (i.e., in conflict with perspective) more often during active trials than during passive trials. A further analysis also revealed that among the parallax responses, there were fewer inverted responses for active trials ($F(1,18) = 40.26$, $p \ll 0.001$). This preference for parallax during active trials as compared to passive trials together with fewer inverted responses on active trials implies a preference for a percept that is consistent with an object that is stationary (i.e., not rotating) in the allocentric frame of reference, as previously reported by Wexler, Panerai et al. (2001). In [Experiments 2](#) and [3](#), we examined whether this stationarity bias can be modulated by experience.

Experiment 2—Stationarity bias modulation

Procedure

The stimuli used in this experiment were the same as the stimuli used in [Experiment 1](#), except that all trials were active, grids comprised 18 rather than 10 squares, and only two values of ΔT were used (0° and 180°). The two values of ΔT corresponded to “training trials” and “test trials,” respectively. Thus, training trials had the no-conflict configuration, with perspective and parallax specifying the same tilt. Additionally, on training trials only (i.e., to prevent duals from being perceived), we added a luminance cue to the grid, such that the luminance of the grid's edges covaried with simulated distance from the observer ([Movie 1](#)). Across 50 cm of depth (25 cm on either side of the display screen), edge luminance decreased by a factor of two. This “proximity-luminance covariance” is a potent cue for depth (Doshier, Sperling, & Wurst, 1986; Schwartz & Sperling, 1983) and provided us

with a better control over observers' percepts on training trials. Two groups of observers (eight in each group) were presented with different types of training trials. One group saw “Moving” stimuli and the other saw “Stationary” stimuli. “Moving” stimuli were constructed by rotating the planar grid contingent on observer motion. The grid was rotated in space about its vertical axis, so as always to have an angular position equal to twice the angle defined by observers' current line of sight (relative to their line of sight at the start of the trial). The perceived motion of the grid in this stimulus mimics the perceived motion on active trials in [Experiment 1](#), when observers perceived the grid to be inverted (see [Footnote 1](#)). “Stationary” stimuli were constructed exactly the same way but without rotation in space. Test trials were identical to stimuli in [Experiment 1](#) with $\Delta T = 180^\circ$ ([Movie 1](#)) and contained no luminance-gradient cues. Therefore, on a given test trial, observers could perceive either a distorted stationary grid (with the simulated tilt) or its rotating regular-shaped dual (with opposite tilt). The task of the observer was to report the perceived tilt that was perceptually coupled (Hochberg & Peterson, 1987) to the apparent rotation of the grid, because duals have both reversed depth and reversed rotation. Hence, the perceived rotation could be discerned from the reported tilt, thus allowing observers to report the grid's apparent rotation indirectly, without knowing that rotation was actually the dependent variable of interest.

Observers ran four blocks of 60 trials each. The first blocks contained only training trials, while the next three blocks contained an equal mix of test and training trials presented pseudo-randomly. The session lasted approximately an hour. As in [Experiment 1](#), observers reported the perceived tilt of the grid using a gauge figure.

Results

[Figure 3](#) shows the histogram of perceived tilts for all observers. Responses are plotted such that a perceived tilt of 0° implies that the grid was perceived as stationary, while 180° implies that it was perceived as rotating. The top and bottom rows represent training and test trials, respectively, while columns represent the two training groups. The responses on training trials (top row) are predominantly near 180° for “Moving” stimuli and 0° for “Stationary” stimuli, which shows that luminance and perspective were effective for determining observers' percepts on training trials.

These data are summarized in [Figure 4](#) using the *Stationarity Preference Index* (SPI), defined as the normalized difference between the proportion of trials on which the stimulus was perceived as stationary and the proportion of trials on which the stimulus was perceived as moving ($SPI = (P_S - P_M)/(P_S + P_M)$), for training and test trials. Responses between -22.5 and $+22.5$ deg were classified as stationary percepts while responses less than

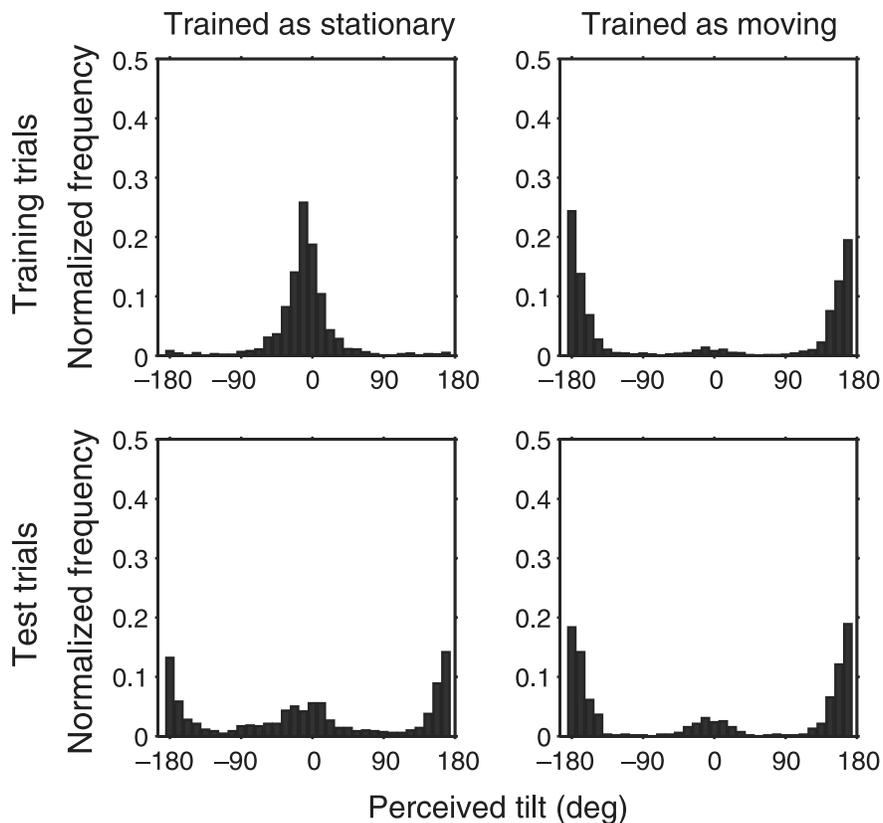


Figure 3. Histograms show the proportion of trials as a function of perceived tilt for all observers. A response of 0° implies that observers perceived the grid as stationary, while a value of 180° implies that observers perceived the grid to be rotating. The top and bottom rows represent training and test trials, respectively, while the columns represent the two training groups pooled across all blocks and observers.

22.5° from 180° were classified as moving percepts. SPI is close to 1 for “Stationary” training trials and close to -1 for “Moving” training trials, showing again that we were able to control observers’ percepts. A two-way ANOVA

with training condition (Moving or Stationary) and block number as factors revealed a highly significant effect of training condition for training trials ($F(1,7) = 3275, p \ll 0.0001$), as expected.

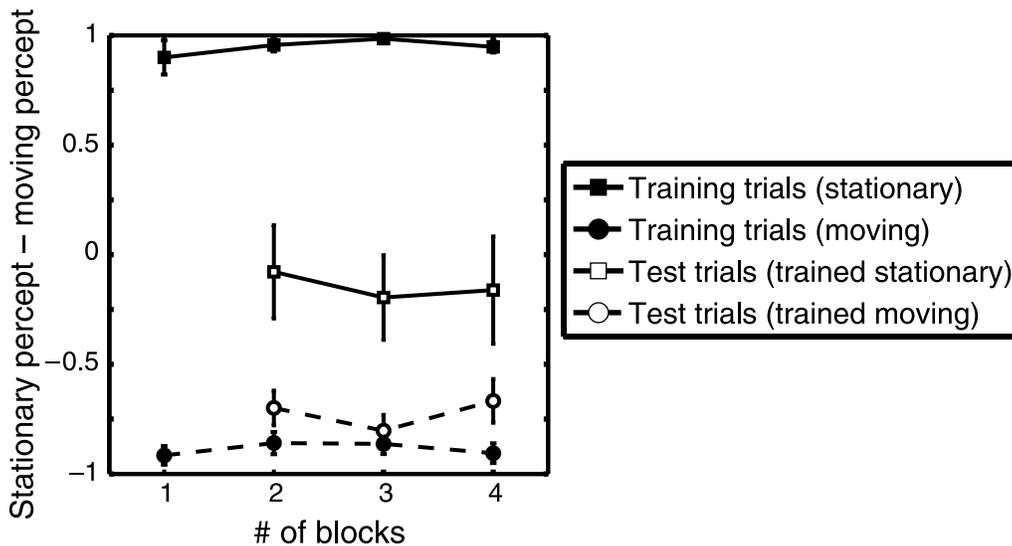


Figure 4. The normalized difference between the proportion of trials on which the stimulus was perceived as stationary and the proportion of trials on which the stimulus was perceived as moving $SPI = (P_S - P_M)/(P_S + P_M)$, for training and test trials.

Our primary interest was whether the training condition affected the appearance of test trials. The effect of training was highly significant ($F(1,7) = 18.36$, $p < 0.001$). The group trained with Stationary stimuli perceived the ambiguous test stimuli as stationary more often than did the group trained with Moving stimuli. There was no effect of block number, nor any significant interaction between the two factors, which suggests that observers' bias was updated fairly quickly to its final value.

The effect of training on the appearance of test trials can also be quantified using a z -score measure (Backus, 2009; Doshier et al., 1986). For this analysis, responses between -90 and $+90$ deg were classified as stationary and the proportion of stationary responses was converted to z -score (using the inverse cumulative normal transformation, i.e., probit). The mean z -score for the group trained on Stationary stimuli was -0.39 (bootstrapped 95% CI $[-0.62 -0.25]$), while the mean z -score for the group trained on Moving stimuli was -1.42 (bootstrapped 95% CI $[-1.65 -1.28]$). Thus, the differential effect of training was approximately 1 normal equivalent deviation.

For test trials, SPI was close to zero for the group that was trained to see the grid as stationary and negative for the group that was trained to see the grid as moving, which might seem at odds with the stationarity bias on first glance. That is not the case. The stimuli used in Experiment 2 were different from stimuli used in Experiment 1 in two critical aspects. First, the number of visible squares in Experiment 2 (18 squares) was nearly twice the number in Experiment 1 (10 squares). This difference should have made linear perspective cues more effective owing to increased square density and longer straight lines. Second, since all corner squares were visible during Experiment 2, relative size was more reliable as a depth cue. These two factors would have increased observers' reliance on perspective cues and hence caused the stimulus to be seen more often as moving in Experiment 2 as compared to Experiment 1.

Experiment 3—Independent biases across visual field

Experiment 3 was designed to determine whether the prior is expressed as a single uniform bias across the visual field or if different priors can be learned, and can operate at different locations in the visual field simultaneously.

Stimuli

The stimuli were constructed in a similar manner as in Experiment 2 with three modifications. All stimuli were half the size (visual angle 10.1°); they were presented at 7.5° above or 7.5° below fixation; and test trial stimuli

were constructed using orthographic projection in order to make the two perceived forms more equally likely (by eliminating the influence of perspective cues, as compared to Experiment 2). Training trials were constructed as in Experiment 2 and contained perspective cues. Again, luminance-proximity covariance was added as a cue to control observers' percept on training trials.

Procedure

We collected data from 16 observers; eight of them were trained to see the grid as moving when it appeared above fixation and as stationary when it appeared below fixation, while the other eight were trained with reverse contingency. Observers ran four blocks of 60 trials each. The first block contained only training trials while the next three contained an equal mix of test and training trials presented pseudo-randomly. Observers were instructed to fixate at the center of the screen while performing head movements. A fixation square was presented to aid fixation. On any given trial, the stimulus was presented at one location (either above or below fixation) followed by the gauge figure at the same location as the grid. The gauge figure was presented after the observer completed the movement cycle.

Results

Figure 5 shows the histogram of perceived tilts for all observers. The response tilts are plotted such that a value of 0° implies that the observers perceived the grid as stationary, while a value of 180° implies that observers' perceived the grid as moving. The top and bottom rows represent training and test trials, respectively, while the columns represent the two types of training.

The data are summarized in Figure 6 using the SPI. We collapsed the data across blocks, as SPI did not vary between blocks either for test or training trials. The left panel shows the data averaged across all observers, the center panel shows data averaged across observers that were trained to see objects above fixation as moving, while the right panel shows data averaged across observers that were trained to see objects above fixation as stationary. A two-way repeated measures ANOVA with location of trained as moving and training as the factors revealed a significant effect of training for both training trials ($F(1,28) = 73.79$, $p \ll 0.0001$) and test trials ($F(1,28) = 6.71$, $p < 0.05$). There was also a significant interaction between the two factors for test trials only ($F(1,28) = 5.83$, $p < 0.05$). The group that was trained with Moving trials above fixation showed a larger effect of location (on the proportion of test trials seen as moving) than did the group that was trained with Moving trials below fixation.

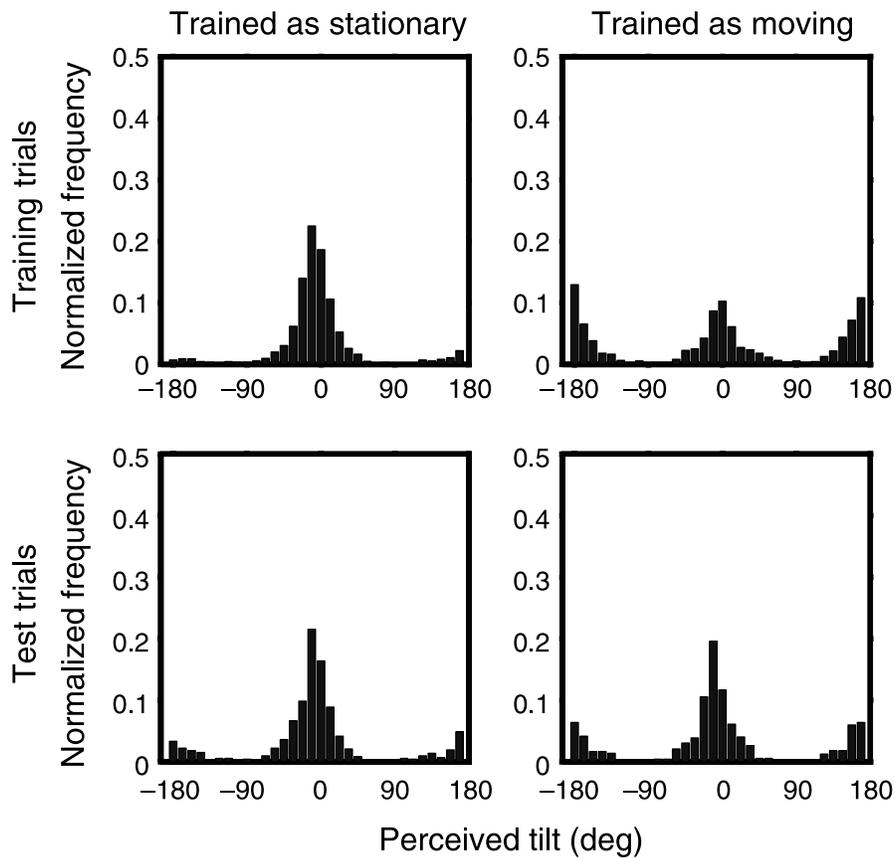


Figure 5. Histograms show the proportion of trials as a function of perceived tilt for all observers. A tilt response of 0° implies that the grid was perceived as stationary, while a response of 180° implies that the grid was perceived as moving. The top and bottom rows represent training and test trials, respectively, while the columns represent the two training types pooled across all observers and all blocks.

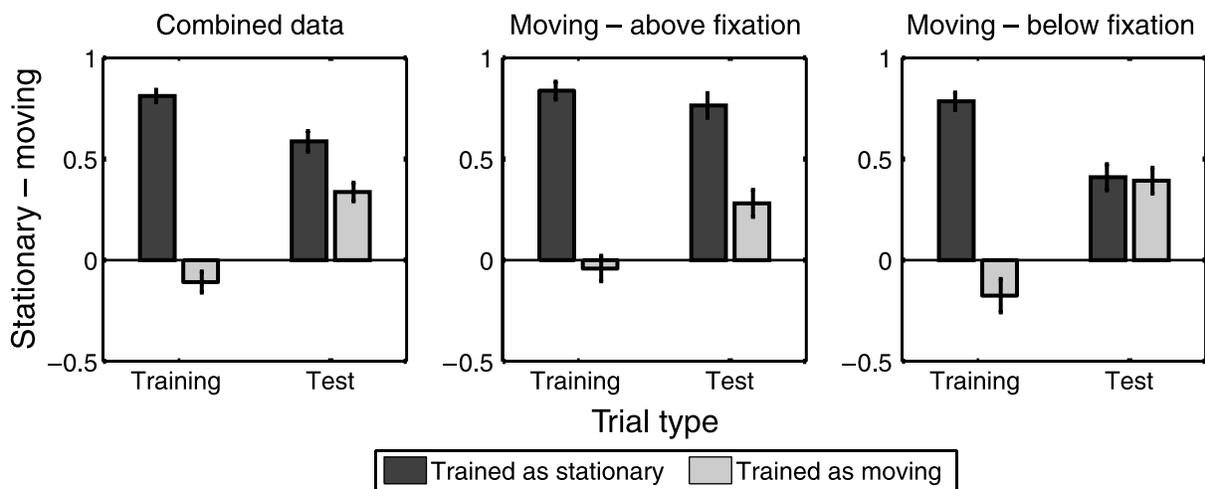


Figure 6. The normalized difference between the proportion of trials on which the stimulus was perceived as stationary and the proportion of trials on which the stimulus was perceived as moving $SPI = (P_S - P_M)/(P_S + P_M)$, for training and test trials. The left panel shows the data averaged across all observers, the center panel shows the data averaged across observers who were trained to see objects above fixation as moving, while the right panel shows the data averaged across observers who were trained to see objects above fixation as stationary.

To quantify the effect of training, we converted the proportion of trials perceived as stationary for both locations to z -scores (Backus, 2009; Doshier et al., 1986) and computed the difference (z -diff measure). The mean z -diff measure averaged across all observers was 0.72 (bootstrapped 95% CI [0.48 1.00]). The mean z -diff measure for observers who were trained with a moving grid above fixation was 1.28 (bootstrapped 95% CI [0.95 1.70]) while the z -diff measure for observers who were trained with a moving grid below fixation was 0.17 (bootstrapped 95% CI [−0.11 0.47]).

It is worth pointing out that luminance and perspective depth cues were not as effective in determining the percept on training trials in the current experiment as they were in Experiment 2. This is a predicted consequence of using smaller stimuli, because perspective cues depend on variation in size across the field of view and are therefore more difficult to measure in small stimuli, while the ratio of luminance from near to far was only half that of Experiment 2 (because the same distance-based rule was used for proximity-luminance covariance). These cues were partially effective but did not completely overcome observers' pre-existing bias to see objects as stationary: observers perceived stationary training trials as stationary most of the time but perceived moving training trials as moving only 50% of the time. Even so, training induced a differential preference for stationarity on ambiguous test trials. Further, as predicted by the stationarity hypothesis (Wexler, Lamouret et al., 2001), observers perceived the ambiguous test trials as stationary more often than moving. Hence, our manipulation to reduce the strength of perspective cues reinstated the stationarity bias that was not observed in Experiment 2 due to the presence of strong perspective and relative size cues. Thus, the only factors that should have affected the perceptual decision on a given trial were the stationarity bias and (unmodeled) decision noise. We discuss the difference between the two groups in this experiment in the General discussion section, below.

General discussion

In Experiment 1, we replicated Wexler, Panerai et al. (2001) finding that the perceptual interpretation of optic flow generated by observer motion favors objects that are stationary in the allocentric frame of reference. We showed that this stationarity bias could be modulated by experience in Experiment 2. Observers were trained to see objects either as moving or as stationary using disambiguating depth cues. After differential training, two groups of observers showed a differential preference that was consistent with training. In Experiment 3, we showed that different biases could be learned by the same observer at different locations in the visual field.

The data from Experiment 1 confirm that the human perceptual system uses extraretinal ego-motion signals to resolve ambiguous optic flow signals. A given optic flow signal was more often perceived as originating from a stationary object than a rotating one, when ego-motion signals were available to disambiguate it. Thus, the human perceptual system has a bias that favors stationary objects in the allocentric frame of reference. It has been argued that this preference results from prior experience: it is reasonable to assume that most optic flows are generated by observer motion relative to stationary objects in the physical world (Wexler, Lamouret et al., 2001). In Bayesian terms, where appearance of the object is understood to manifest the system's belief about the world, observers behaved as though they have a prior belief that objects are more likely to be stationary than rotating. This prior is presumably a reflection of the environmental statistics. This stationarity prior can be seen as analogous to the "slowness" prior (Weiss & Adelson, 1998; Weiss et al., 2002) used to explain the perceived speeds of low contrast stimuli. Weiss et al. (2002) argued that slow motions are more frequent than faster ones, and that an ideal observer should exploit this fact, in which case low contrast stimuli would be expected to have lower apparent speed than high contrast stimuli. Freeman, Champion, and Warren (2010) have also used the slowness prior to explain various motion illusions observed under smooth pursuit eye movements. The "slowness" prior describes a distribution of retinal motions with a peak at zero; analogously, the preference for 3D stationarity observed here and in previous studies (Wexler, Panerai et al., 2001) would presumably describe a distribution of object angular velocities in the allocentric frame of reference that peaks at zero. A prior for 3D object motion is presumably more complex than for 2D retinal motion and might take context into account in various ways (whether the observer is moving, whether the object is attached to the ground, and so on).

In Experiment 2, we examined whether the stationarity prior can be modified by altering the environmental statistics in short (1 h) training sessions. The observers' preference for stationarity was modulated in the direction consistent with training. The group that was trained with moving stimuli perceived the ambiguous test stimuli as moving more often than the group that was trained with stationary stimuli. This result adds to an accumulating literature in support of the view that experience affects how we perceive the world, even into adulthood (Flanagan, Bittner, & Johansson, 2008; Haijiang et al., 2006). The neural mechanisms that implement this learning might or might not represent the stationarity prior explicitly. Thus, one possibility is that training directly altered a mechanism that utilizes ego-motion signals to interpret optic flow signals; in the absence of optic flow, no representation of the stationarity prior would exist in the brain. A second possibility is that a module that combines optic flow signals with ego-motion signals represents both possible interpre-

tations simultaneously, with the preference being expressed at a decision stage after the prior for stationarity is taken into account. Our data are consistent with either of these accounts; they show that computationally there was an effective change in the stationarity prior, but they do not provide a means of distinguishing which class of neural mechanisms or representations are responsible.

Other studies have shown a change in the interpretation of signals after training (Adams et al., 2004; Champion & Adams, 2007; Jacobs & Fine, 1999; Sinha & Poggio, 1996), but they did not utilize multiple perceptual signals. Adams et al. (2004) show that the light-from-above prior is modified after training in a shape-from-shading task so that shape-from-shading cues were interpreted differently. In this case, the shading cue is the only perceptual signal. Similarly, Sinha and Poggio (1996) showed that objects selected from a pre-exposed training set were less likely to be perceived as non-rigid compared to objects that were novel in a structure-from-motion task. The rigidity assumption was modulated in their experiments just like the stationarity assumption was modulated in our experiments. Our results extend the two findings mentioned above. The human perceptual system can learn not only to combine arbitrary signals and learn novel meanings for a signal but can also learn to change how it behaves when integrating information from two signals from different sensory modalities (here, retinal and non-retinal). Our results are also analogous to those observed by Haarmerier, Bunjes, Lindner, Berret, and Thier (2001) who also showed that the perception of stability during eye movements can be modulated by experience. Tcheang, Gilson, and Glennerster (2005) also described a drift in bias for stability/stationarity after observers had been exposed to rotating stimuli; however, they did not measure the drift explicitly.

In [Experiment 3](#), observers showed a differential bias for stationarity between two different locations that was consistent with the training after exposure to moving stimuli above fixation and stationary stimuli below fixation (or vice versa) during training. In other words, location of the stimuli was recruited as a cue (Backus & Haijiang, 2007; Haijiang et al., 2006) to affect the interpretation of optic flow signals. The current experimental set-up cannot distinguish between retinotopic learning and learning in allocentric frame of reference, i.e., whether the stationarity prior was modulated at different retinotopic locations or at different locations in the world. A great deal of perceptual learning occurs in retinotopic coordinates (Backus & Haijiang, 2007; Harrison & Backus, 2010a; Sagi & Tanne, 1994). However, the fact that the human perceptual system prefers the percept that is most consistent with a stationary object in allocentric frame of reference implies that it has access to it. Recent electrophysiological studies have shown that neurons with allocentric properties exist in the human brain (Zaehle et al., 2007) and that there are place cells in human hippocampus

(Ekstrom et al., 2003). Studies involving infants have shown that immediately after the infants learn to crawl they become more adept at using visual information (optic flow) for posture control compared to infants of similar age who are yet to learn crawling (Higgins, Campos, & Kermoian, 1996). Locomotion also assists in the development of allocentric spatial maps (Bai & Bertenthal, 1992). Therefore, signals needed to instantiate a preference for stationarity are present in the visual system.

We also found in [Experiment 3](#) that learning was apparently stronger when objects above fixation were trained as moving, and objects below fixation were trained as stationary, as compared to the case when the association was reversed. Why might that have been the case? We see two plausible explanations. First, observers could have started the experiment with equal bias to perceive objects as stationary at both locations. In that case, learning must have occurred only in the group that was trained on moving stimuli above fixation. Alternatively, the perceptual system might have a stronger pre-existing bias to perceive objects as stationary in the lower visual field than it does in the upper visual field. In that case, the effect of training was to diminish the difference in bias in one case (when objects above fixation were trained as stationary) and to enhance it in the other case (when objects above fixation were trained as moving). This explanation is plausibly rooted in the statistics of natural scenes; objects in the lower visual field might more often be rigidly attached to the ground, for example. The data from test trials are consistent with either hypothesis.

The magnitude of the effects we measured here were smaller than many effects observed previously (Haijiang et al., 2006; Harrison & Backus, 2010a, 2010b), which were an order of magnitude larger when expressed in z-score units (or normal equivalent deviations). It is not surprising, however, that very different types of prior belief would be learned at different rates, especially when effect sizes are in units of standard deviations in decision noise for these rather different types of dichotomous perceptual decisions. The key difference between the studies mentioned above that observed strong learning and the current study is in the frame of reference in which the learning occurred. In previous studies, learning occurred as well as manifested itself in retinotopic frame of reference (Harrison & Backus, 2010a), while in our study, even though it is not clear whether the learning occurred in retinotopic or in allocentric frame of reference, it was manifested in the allocentric frame of reference. It is possible that due to the retinotopic organization of the early visual system, retinotopic location is privileged and retinotopic location contingent biases are more strongly learnt. Indeed, learning is weaker for signals other than retinotopic location such as translation direction (Haijiang et al., 2006) or vertical disparity (Di Luca, Ernst, & Backus, 2010).

In our experiments, observers made voluntary movements that generated ego-motion and the resulting optic flow. It would be interesting to explore whether similar results are obtained when observers are moved using external force. If so, such a scenario would suggest the involvement of a system that passively measures ego-motion, like the vestibular system. There is some experimental evidence (Nadler, Nawrot, Angelaki, & DeAngelis, 2009; Nawrot, 2003) that the visual system uses the slow optokinetic response signal in conjunction with the translational vestibular ocular response to disambiguate optic flow signals generated by ego-motion. A failure to observe similar findings during passive locomotion would suggest that the efference copy of the ego-motion signal is utilized to interpret retinal optic flow. Recent studies have shown that activity in place cells in monkeys (Foster, Castro, & McNaughton, 1989; Nishijo, Ono, Eifuku, & Tamura, 1997) and rats (Song, Kim, Kim, & Jung, 2005) and head direction cells in rats (Stackman, Golob, Bassett, & Taube, 2003; Zugaro, Tabuchi, Fouquier, Berthoz, & Wiener, 2001) is reduced during passive ego-motion as compared to active ego-motion. One might therefore predict a stronger stationarity bias for active ego-motion. Indeed, Wexler (2003) showed that the stationarity prior is weakened when there is a mismatch between observer motion and motor command.

In summary, our results add to a short but growing list of demonstrated cases wherein experience affects appearance. The human perceptual system modulates its utilization of extraretinal ego-motion signals in conjunction with optic flow signals, based on environmental statistics. Moreover, this modulation can occur independently at two separate locations in the visual field. This modulation may best be conceived as an experience-dependent change in a Bayesian prior for object stationarity during ego-motion. Further experiments are required to determine the neural correlates, retinotopicity, and longevity of this learning.

Acknowledgments

We thank Benjamin Hunter McFadden for assisting with data collection and Martha Lain for assisting with recruitment of observers. This study was supported by HFSP Grant RPG 2006/3, NSF Grant BCS-0617422, and NIH Grant EY-013988 to B. T. Backus.

Commercial relationships: none.

Corresponding author: Dr. Anshul Jain.

Email: ajain@sunyopt.edu.

Address: Graduate Center for Vision Research, SUNY College of Optometry, 33 W. 42nd Street, New York, NY 10036, USA.

Footnote

¹To understand why the dual of the stationary object rotates (in the world) by twice the observer's change in viewing direction, consider what the observer sees as he or she fixates the object. The stationary object creates a pattern of flow within the retinal image that is due to relative rotation between the observer and the object. The depth-inverted dual will create the same retinal flow if and only if it rotates by the same amount in the opposite direction *relative to the observer*. Thus, if the observer's angle of regard (relative to the room) rotates by θ deg as he or she moves (while fixating the object), then, assuming orthographic projection, his or her retinal images could equally have been created, during the same self-motion, by the dual as it rotates through 2θ deg.

References

- Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience*, 7, 1057–1058.
- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257–262.
- Backus, B. T. (2009). The mixture of Bernoulli experts: A theory to quantify reliance on cues in dichotomous perceptual decisions. *Journal of Vision*, 9(1):6, 1–19, <http://www.journalofvision.org/content/9/1/6>, doi:10.1167/9.1.6. [PubMed] [Article]
- Backus, B. T., & Haijiang, Q. (2007). Competition between newly recruited and pre-existing visual cues during the construction of visual appearance. *Vision Research*, 47, 919–924.
- Bai, D. L., & Bertenthal, B. I. (1992). Locomotor status and the development of spatial search skills. *Child Development*, 63, 215–226.
- Carter, O., & Cavanagh, P. (2007). Onset rivalry: Brief presentation isolates an early independent phase of perceptual competition. *PLoS One*, 2, e343.
- Champion, R. A., & Adams, W. J. (2007). Modification of the convexity prior but not the light-from-above prior in visual search with shaded objects. *Journal of Vision*, 7(13):10, 1–10, <http://www.journalofvision.org/content/7/13/10>, doi:10.1167/7.13.10. [PubMed] [Article]
- Di Luca, M., Ernst, M. O., & Backus, B. T. (2010). Learning to use an invisible visual signal for perception. *Current Biology*, 20, 1860–1863.

- Domini, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: A new theory. *Trends in Cognitive Sciences*, 7, 444–449.
- Dosher, B. A., Sperling, G., & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research*, 26, 973–990.
- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., et al. (2003). Cellular networks underlying human spatial navigation. *Nature*, 425, 184–188.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429–433.
- Ernst, M. O., Banks, M. S., & Bulthoff, H. H. (2000). Touch can change visual slant perception. *Nature Neuroscience*, 3, 69–73.
- Fantoni, C., Caudek, C., & Domini, F. (2010). Systematic distortions of perceived planar surface motion in active vision. *Journal of Vision*, 10(5):12, 1–20, <http://www.journalofvision.org/content/10/5/12>, doi:10.1167/10.5.12. [PubMed] [Article]
- Flanagan, J. R., Bittner, J. P., & Johansson, R. S. (2008). Experience can change distinct size–weight priors engaged in lifting objects and judging their weights. *Current Biology*, 18, 1742–1747.
- Foster, T. C., Castro, C. A., & McNaughton, B. L. (1989). Spatial selectivity of rat hippocampal neurons: Dependence on preparedness for movement. *Science*, 244, 1580–1582.
- Freeman, T. C., Champion, R. A., & Warren, P. A. (2010). A Bayesian model of perceived head-centered velocity during smooth pursuit eye movement. *Current Biology*, 20, 757–762.
- Haarmeier, T., Bunjes, F., Lindner, A., Berret, E., & Thier, P. (2001). Optimizing visual motion perception during eye movements. *Neuron*, 32, 527–535.
- Haijiang, Q., Saunders, J. A., Stone, R. W., & Backus, B. T. (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 483–488.
- Harrison, S. J., & Backus, B. T. (2010a). Disambiguating Necker cube rotation using a location cue: What types of spatial location signal can the visual system learn? *Journal of Vision*, 10(6):23, 1–15, <http://www.journalofvision.org/content/10/6/23>, doi:10.1167/10.6.23. [PubMed] [Article]
- Harrison, S. J., & Backus, B. T. (2010b). Uninformative visual experience establishes long term perceptual bias. *Vision Research*, 50, 1905–1911.
- Helmholtz, H. v. (1910/1925). *Treatise on physiological optics* (vol. III) (J. P. C. S. o. G. 1910, Trans.). New York: Optical Society of America.
- Higgins, C. I., Campos, J. J., & Kermoian, R. (1996). Effect of self-produced locomotion on infant postural compensation to optic flow. *Developmental Psychology*, 32, 836–841.
- Hochberg, J., & Peterson, M. A. (1987). Piecemeal organization and cognitive components in object perception: Perceptually coupled responses to moving objects. *Journal of Experimental Psychology: General*, 116, 370–380.
- Jacobs, R. A., & Fine, I. (1999). Experience-dependent integration of texture and motion cues to depth. *Vision Research*, 39, 4062–4075.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, 55, 271–304.
- Knill, D. C. (2007). Learning Bayesian priors for depth perception. *Journal of Vision*, 7(8):13, 1–20, <http://www.journalofvision.org/content/7/8/13>, doi:10.1167/7.8.13. [PubMed] [Article]
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, 26, 161–179.
- Koenderink, J. J., & van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22, 773–791.
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, 35, 389–412.
- Langer, M. S., & Bulthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception*, 30, 403–410.
- Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B: Biological Sciences*, 208, 385–397.
- Nadler, J. W., Nawrot, M., Angelaki, D. E., & DeAngelis, G. C. (2009). MT neurons combine visual motion with a smooth eye movement signal to code depth-sign from motion parallax. *Neuron*, 63, 523–532.
- Nawrot, M. (2003). Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research*, 43, 1553–1562.
- Nishijo, H., Ono, T., Eifuku, S., & Tamura, R. (1997). The relationship between monkey hippocampus place-related neural activity and action in space. *Neuroscience Letters*, 226, 57–60.
- Papathomas, T. V., & Bono, L. M. (2004). Experiments with a hollow mask and a reverspective: Top-down

- influences in the inversion effect for 3-D stimuli. *Perception*, *33*, 1129–1138.
- Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125–134.
- Rogers, B., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, *22*, 261–270.
- Sagi, D., & Tanne, D. (1994). Perceptual learning: Learning to see. *Current Opinion in Neurobiology*, *4*, 195–199.
- Schwartz, B. J., & Sperling, G. (1983). Luminance controls the perceived 3D structure of dynamic 2D displays. *Bulletin of the Psychonomic Society*, *17*, 456–458.
- Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception [letter]. *Nature*, *385*, 308.
- Sinha, P., & Poggio, T. (1996). Role of learning in three-dimensional form perception. *Nature*, *384*, 460–463.
- Song, E. Y., Kim, Y. B., Kim, Y. H., & Jung, M. W. (2005). Role of active movement in place-specific firing of hippocampal neurons. *Hippocampus*, *15*, 8–17.
- Sperling, G., Doshier, B. A., & Landy, M. S. (1990). How to study the kinetic depth effect experimentally. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 445–450.
- Stackman, R. W., Golob, E. J., Bassett, J. P., & Taube, J. S. (2003). Passive transport disrupts directional path integration by rat head direction cells. *Journal of Neurophysiology*, *90*, 2862–2874.
- Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience*, *1*, 183–184.
- Tcheang, L., Gilson, S. J., & Glennerster, A. (2005). Systematic distortions of perceptual stability investigated using immersive virtual reality. *Vision Research*, *45*, 2177–2189.
- van Boxtel, J. J., Wexler, M., & Droulez, J. (2003). Perception of plane orientation from self-generated and passively observed optic flow. *Journal of Vision*, *3*(5):1, 318–332, <http://www.journalofvision.org/content/3/5/1>, doi:10.1167/3.5.1. [[PubMed](#)] [[Article](#)]
- van Ee, R., Adams, W. J., & Mamassian, P. (2003). Bayesian modeling of cue interaction: Bistability in stereoscopic slant perception. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *20*, 1398–1406.
- Wallach, H., & O’Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205–217.
- Wallach, H., Stanton, L., & Becker, D. (1974). The compensation for movement-produced changes of object orientation. *Perception & Psychophysics*, *15*, 339–343.
- Watanabe, K., & Shimojo, S. (2001). When sound affects vision: Effects of auditory grouping on visual motion perception. *Psychological Science*, *12*, 109–116.
- Weiss, Y., & Adelson, E. H. (1998). *Slow and smooth: A Bayesian theory for the combination of local motion signals in human vision* (No. AI Memo No. 1624; CBCL Paper No. 158). Cambridge, MA: MIT.
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*, 598–604.
- Wexler, M. (2003). Voluntary head movement and allocentric perception of space. *Psychology Science*, *14*, 340–346.
- Wexler, M., Lamouret, I., & Droulez, J. (2001). The stationarity hypothesis: An allocentric criterion in visual perception. *Vision Research*, *41*, 3023–3037.
- Wexler, M., Panerai, F., Lamouret, I., & Droulez, J. (2001). Self-motion and the perception of stationary objects. *Nature*, *409*, 85–88.
- Zaehle, T., Jordan, K., Wustenberg, T., Baudewig, J., Dechent, P., & Mast, F. W. (2007). The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, *1137*, 92–103.
- Zugaro, M. B., Tabuchi, E., Fouquier, C., Berthoz, A., & Wiener, S. I. (2001). Active locomotion increases peak firing rates of anterodorsal thalamic head direction cells. *Journal of Neurophysiology*, *86*, 692–702.