

# Perception of object trajectory: Parsing retinal motion into self and object movement components

**Paul A. Warren**

School of Psychology, Tower Building,  
Cardiff University Cardiff, Wales, UK



**Simon K. Rushton**

School of Psychology, Tower Building,  
Cardiff University Cardiff, Wales, UK



A moving observer needs to be able to estimate the trajectory of other objects moving in the scene. Without the ability to do so, it would be difficult to avoid obstacles or catch a ball. We hypothesized that neural mechanisms sensitive to the patterns of motion generated on the retina during self-movement (optic flow) play a key role in this process, “parsing” motion due to self-movement from that due to object movement. We investigated this “flow parsing” hypothesis by measuring the perceived trajectory of a moving probe placed within a flow field that was consistent with movement of the observer. In the first experiment, the flow field was consistent with an eye rotation; in the second experiment, it was consistent with a lateral translation of the eyes. We manipulated the distance of the probe in both experiments and assessed the consequences. As predicted by the flow parsing hypothesis, manipulating the distance of the probe had differing effects on the perceived trajectory of the probe in the two experiments. The results were consistent with the scene geometry and the type of simulated self-movement. In a third experiment, we explored the contribution of local and global motion processing to the results of the first two experiments. The data suggest that the parsing process involves global motion processing, not just local motion contrast. The findings of this study support a role for optic flow processing in the perception of object movement during self-movement.

**Keywords:** optic flow, perceived trajectory, object movement, self-movement, induced motion, flow parsing

**Citation:** Warren, P. A., & Rushton, S. K. (2007). Perception of object trajectory: Parsing retinal motion into self and object movement components. *Journal of Vision*, 7(11):2, 1–11, <http://journalofvision.org/7/11/2/>, doi:10.1167/7.11.2.

## Introduction

Motion of the image of an object on the retina indicates that object position relative to the observer has changed, but it does not indicate whether this is due to object movement, observer movement, or some combination of the two. As a consequence, determining whether or how an object has moved within the scene requires information about both object position and self-movement, so that the retinal motion arising from self-movement can be factored out (Wallach, 1987).

Under many circumstances, it should be possible to estimate self-movement from “extra-retinal” information (i.e., vestibular and proprioceptive information, efferent motor signals). Perception of object movement when extra-retinal information is available has been extensively studied by Gogel (1990), Swanston and Wade (1988), Swanston, Wade, Ono, and Shibuta (1992), Wallach (1987), and more recently by Tcheang, Gilson, and Glennerster (2005) and van Pelt and Medendorp (2007).

However, there are situations in which accurate extra-retinal information is not available. For example, when moving at a constant velocity the vestibular system provides little information. Similarly, when traveling in a car, the efferent motor signals contain little information about observer movement.

Furthermore, the brain appears to have the functionality required to solve this problem from visual or retinal information alone. Extensive psychophysical research has demonstrated the ability of the primate brain to identify the structured patterns of global retinal motion that are characteristic of self-movement (optic flow). Neurophysiological and imaging research have identified areas MT/V5 and MSTd, among others, as candidate neural substrates for this sensitivity (for a review, see Lappe, Bremmer, & van den Berg, 1999).

Optic flow processing is behaviorally important for the estimation of the movement of the eyes in space (e.g., Lappe et al., 1999; Warren & Hannon, 1988), for the control of posture (e.g., Lee & Aronson, 1974), for updating memory of egocentric position (e.g., Brecher, Brecher, Kommerell, Sauter, & Sellerbeck, 1972; Lepecq, Jouen, & Dubon, 1993), and for the stabilization of gaze during self-movement (e.g., Bussetini, Masso, & Miles, 1997; Miles, 1997; Zhou, Wei, & Angelaki, 2002). The hypothesis we have recently proposed (Rushton, Bradshaw, & Warren, 2006; Rushton & Warren, 2005; see also Royden, Wolfe, & Klempen, 2001) and examined further here is that optic flow processing also underpins our ability to estimate the trajectory of a moving object during self-movement.

We suggest that optic flow detectors act as filters, parsing retinal flow into distinct components due to observer and object movement. The component of retinal

motion due to self-movement can then be “subtracted” from the total retinal flow, leaving only those components of motion due to movement of objects in the world. As a consequence, movement of an object relative to the background scene can be estimated as if the observer was stationary. This idea can be compared to Johansson’s (1974) vector analysis proposal. Johansson suggested that the visual system decomposes a pattern of retinal motion into a common 2D component of motion and a relative motion component. In Johansson’s terms, we are proposing that retinal motion is decomposed into a common 3D component of motion and a relative motion component.

In a previous study, we have demonstrated results compatible with the use of “flow parsing” in the detection of movement of a target object relative to a rigid scene during self-movement (Rushton & Warren, 2005). In this study, the task we set observers is more difficult; in a complex pattern of retinal motion, we ask for an estimate of the trajectory of a target (probe) object.

To demonstrate that trajectory estimation is indeed dependent upon optic flow processing, we exploit a difference between two flow fields arising from different observer movements. When the eye rotates, the associated pattern of retinal flow is independent of the distance of objects in the scene. However, when the eye translates laterally, the associated pattern of retinal flow is dependent on object distance; near objects move faster and further than far objects. Consequently, if the component of retinal flow arising from self-movement is recognized and parsed from the retinal image, the parsed component should be independent of depth in the case of observer rotation but dependent on depth in the case of observer translation.

To illustrate this idea, imagine an observer making a leftward eye or (to a first approximation) head rotation. Consider a target that maintains its position directly ahead of the observer (left half of Figure 1A). Leftward rotation produces a uniform pattern of rightward motion,  $h$ , across the visual field. If the brain parses out self-movement flow

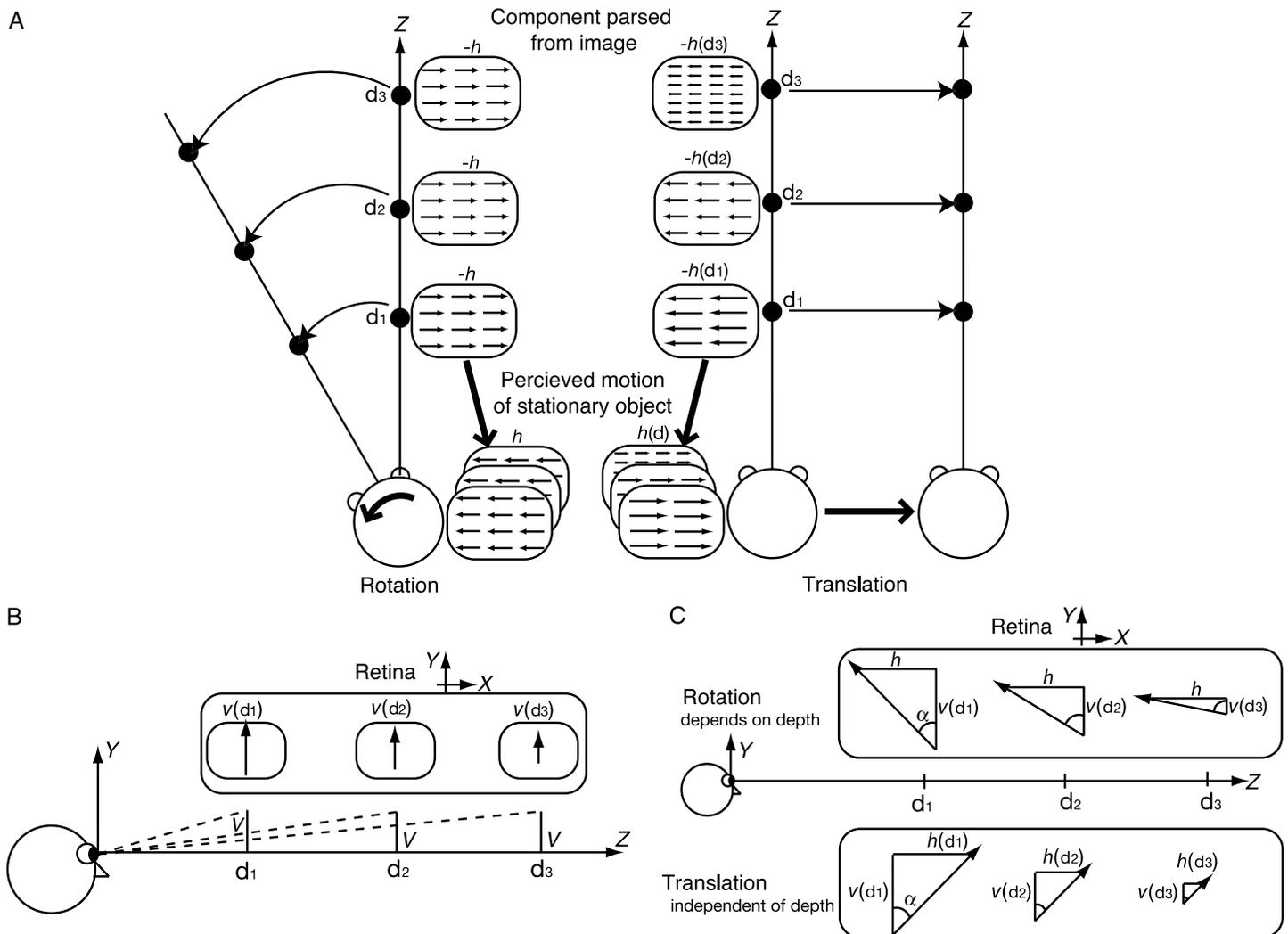


Figure 1. Predictions of perceived probe motion during different self-movements under the flow parsing hypothesis. (A) Shows predicted perceived probe motion for horizontal rotation and translation. (B) The retinal motion of a probe when vertical physical movement scales with distance. (C) Combined vertical and horizontal perceived motion components from physical and perceived probe movements.

components, it will subtract the same rightward motion component from the retinal motion of every object in the visual field, irrespective of distance to the object. The subtraction of a rightward motion component from a stationary retinal target image will effectively add a leftward target motion component that is *independent* of probe depth. Now consider an observer making a rightward translation while the target probe remains directly ahead (right half of [Figure 1A](#)). Rightward translation produces a nonuniform pattern of leftward motion,  $h$ , on the retina that varies with depth. If the brain parses self-movement flow components, it will now subtract a leftward component of motion from objects in the visual field; however, in this case, the subtracted component will be dependent on probe depth. Subtraction of such a leftward motion component from a stationary retinal target image will effectively add a rightward target motion component that is *dependent* on probe depth.

We simulated similar situations to those described above in an experimental setting. Instead of moving the observers, they were kept stationary and presented with optic flow stimuli (containing stereoscopic disparity) consistent with different types of self-movement. We measured the effect of presentation of such optic flow stimuli on perceived trajectory of a probe object placed at different depths in the visual scene and looked for consistency with the predictions of the flow parsing hypothesis. In a series of psychophysical studies, we demonstrate results that are in line with these predictions and thus consistent with the flow parsing hypothesis.

## General methods

### Observers

Six observers participated in [Experiments 1 and 2](#) and four observers participated in [Experiment 3](#). All were staff

or students within the School of Psychology at Cardiff University. Two of the observers were authors; the rest of the observers were naive as to the purpose of the experiment. All observers had normal or corrected-to-normal vision; none had been found to have any deficits in perception of stereo depth during previous experimental work. Observers' participation in the experimental studies was regulated by the Ethics Committee of the School of Psychology, Cardiff University.

### Stimulus and displays

In the three experiments reported, we generated patterns of retinal motion through simulation of self-movement. At all times the observer remained stationary (with the head on a chin rest).

Twenty-four textured cubes of approximately  $2 \times 2 \times 2$  cm were randomly oriented and were placed within a volume of  $26 \times 26 \times 50$  cm (see [Figure 2](#)), with the center of the volume at either 105 cm ([Experiments 1 and 2](#)) or 115 cm ([Experiment 3](#)) directly ahead of the observer. The observer viewed the objects displayed on a 22-in. Viewsonic p225F CRT in a pitch-black room. Images were rendered using OpenGL and were antialiased. Objects were drawn in red (because it has the fastest phosphor) and a red filter was placed in front of the screen (to increase contrast). Left and right eye views were each generated at 50 Hz, temporally interleaved and viewed through synchronized shutter glasses (Stereographics, CA). The result was a display that produced a compelling percept of 3D objects floating in space. During the experiment, the cubes were moved and transformed in a manner consistent with either an eye rotation of 0.75 deg/s or a lateral eye translation of 2 cm/s. Due to the richness of the displays used, observers perceived distinctly different patterns of movement between the scene and the observer in these two cases.

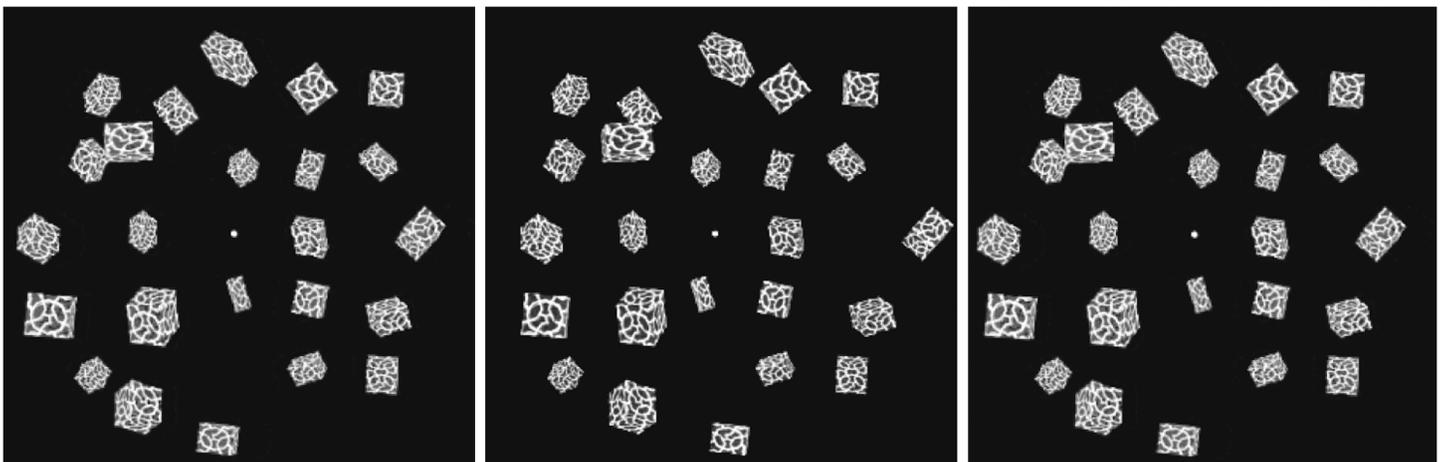


Figure 2. Stereo pairs showing an example of a single frame of the cube array stimulus and the probe dot at middle depth. The left pair of images can be cross-fused to reveal stereo information. The right pair of images can be uncross-fused to show the same image.

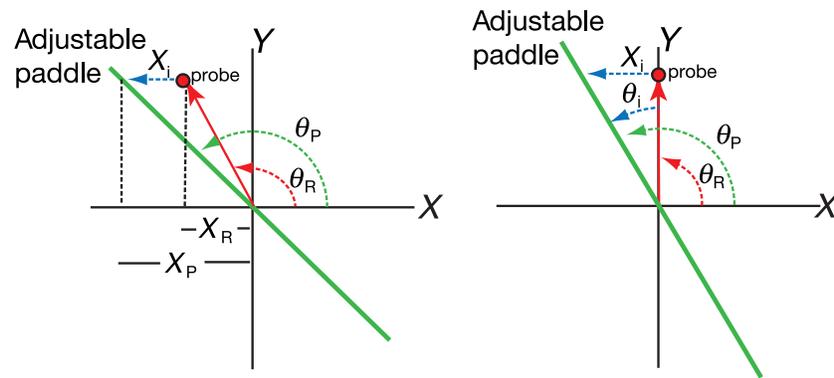


Figure 3. Schematic diagram of the axes and the paddle adjusted by the observers. For illustration only, coordinate system and key parameters are also shown. Absolute angles are measured relative to the X-axis in a counterclockwise direction. Quantities  $X_i$  and  $\theta_i$  characterize the perceived change in trajectory. Left panel shows case when probe does not travel vertically. Right panel shows the special case in which the probe moves vertically (i.e.,  $\theta_R = \pi/2$ ).

A probe dot was simultaneously presented with the cubes and moved in one of five linear directions, which varied by  $0^\circ$ ,  $\pm 15^\circ$ , or  $\pm 30^\circ$  with respect to vertical in a frontoparallel plane at one of three distances. The distances of these planes were defined relative to the center of the volume of the cubes, which coincided with the viewing distance (105 cm in Experiments 1 and 2 and 115 cm in Experiment 3). In Experiments 1 and 2, the planes were offset by  $-20$ ,  $0$ , and  $+20$  cm from the fixation distance, and in Experiment 3 they were offset by  $-18$ ,  $0$ , and  $+18$  cm. The probe moved at a constant speed of  $0.9$  cm/s in the world ( $0.49$  and  $0.45$  deg/s at the viewing distances of  $105$  cm and  $115$  cm in the first two and final experiment, respectively).

## Procedure

In all experiments, at the beginning of each trial observers saw a small fixation cross in the center of the screen for  $0.75$  s (75 frames). Subsequently, observers saw the array of cubes and the probe initially positioned at the center of the screen. The cubes and probe remained stationary for  $0.5$  s (50 frames) so that observers had time to fuse the stereoscopic images. The cubes in the array and the probe dot then moved for  $2$  s (200 frames) and the observers were instructed to fixate the moving probe dot.<sup>1</sup>

Using a modified version of the tilt test (Heckmann & Howard, 1991; Howard & Childerson, 1994), observers were then asked to report the perceived 2D (in a frontoparallel plane) trajectory of the probe. Immediately after presentation of the stimulus, the observer saw on-screen axes (a cross defined by horizontal and vertical lines) with an adjustable “paddle” passing through the origin (see Figure 3). To change the angle of the paddle, observers turned a knob linked to a linear potentiometer. The axes and the adjustable paddle were always presented in the plane of the screen. Any difference in the trajectory

indicated by the observer and the physical trajectory of the probe was assumed to be due to the optic flow manipulation.

## Analysis

For all three experiments reported, observer responses are coded as an angular measure  $\theta_i$ . This quantity is derived from the illusory or the “induced” horizontal motion seen by observers but is deliberately converted back to an angular quantity because, in this experiment, observers made angular adjustments (the results are similar when the horizontal motion is considered). The coordinate system and sign conventions used are shown in Figure 3.

Angles describing “real” physical probe trajectory,  $\theta_R$ , and perceived trajectory,  $\theta_P$ , are defined relative to the X-axis (counterclockwise +ve).

It was assumed that the simulated horizontal observer movement did not interfere with the perceived vertical motion of the probe (i.e., the perceived and the physical vertical motion were identical). Consequently, any difference between the perceived and the physical angles was due to an additional perceived horizontal component of probe motion  $X_i$ . Using simple trigonometry, given the distance traveled by the probe,  $r$ , this quantity can be calculated as

$$X_i = X_R - X_P = r \sin(\theta_P - \theta_R) / \sin(\theta_P). \quad (1)$$

For the sake of consistency with angular measures (which increase in a counterclockwise direction),  $X_i$  is defined as positive to the left.

To make the measurement commensurate across the different probe direction conditions, we calculated it as follows. The component  $X_i$  was transformed back to an

angular quantity  $\theta_i$ , as if it had been induced relative to a physical probe motion that was purely vertical:

$$\theta_i = \tan^{-1}(X_i/r). \quad (2)$$

This quantity also increases in a counterclockwise direction but is now measured relative to the  $Y$ -axis. Note that for the condition in which the physical probe trajectory really is vertical ( $\theta_R = \pi/2$ ), Equations 1 and 2 reduce to  $\theta_i = \theta_P - \pi/2$ , as required (see Figure 3, right panel). In the results that follow,  $\theta_i$  is referred to as the induced tilt.

## Experiments 1 and 2

In the first two experiments, observers saw the cube array and the probe dot described above. In Experiment 1, the cube array moved in a manner consistent with rotation of the observer's eye. In Experiment 2, the cubes moved in a manner consistent with lateral translation of the observer's eye.

Let us revisit the predictions of the flow parsing hypothesis discussed in the Introduction section and consider the implications of these predictions for our experiments. Consider the situation in which the physical motion of the probe is vertical. In this condition, in both experiments, the probe remains directly ahead of the observer and moves vertically upward at the same speed, irrespective of depth. Consequently, due to perspective projection, the vertical retinal motion of the probe,  $v$ , resulting from its movement in the scene,  $V$ , is dependent on probe distance (Figure 1B).

In Experiment 1, we simulate observer rotation. Under the flow parsing hypothesis, the perceived horizontal

motion,  $h$ , of the probe should be independent of probe distance (left half of Figure 1A). In Experiment 2, we simulate observer translation. Under the flow parsing hypothesis, the perceived horizontal motion,  $h$ , of the probe should now be dependent on probe distance (right half of Figure 1A).

Assuming the vectorial combination of (i) perceived motion components due to physical probe movement and flow parsing and (ii) horizontal and vertical perceived motion components, the perceived target trajectory angle should then be dependent upon distance in Experiment 1 and independent of distance in Experiment 2 (Figure 1C).

## Results and discussion

The results from the first experiment (simulated rotation) are shown in Figure 4 for six observers (circles). The results have been averaged over leftward and rightward self-movement conditions and, to preserve agreement with the sign convention, are reported as if the observer always undertook a leftward self-movement (and hence the induced tilt is +ve counterclockwise). At all three probe distances, observers perceived a pronounced motion in the direction of self-movement—or opposite the direction of movement of the background elements if we are to compare this to the classical induced motion effect (Duncker, 1929). As predicted by the flow parsing hypothesis, the induced tilt increased with probe distance.

In the second experiment (simulated translation), the observer judgments depend less on probe distance (Figure 4, left, squares). This is particularly the case for three observers (C.X.B., B.A.E., and S.K.R.) whose trajectory judgments show little dependence on probe distance. Mean judgments averaged over all observers are shown in Figure 4, right. In addition, the figure shows the induced tilts obtained for a model implementing flow parsing

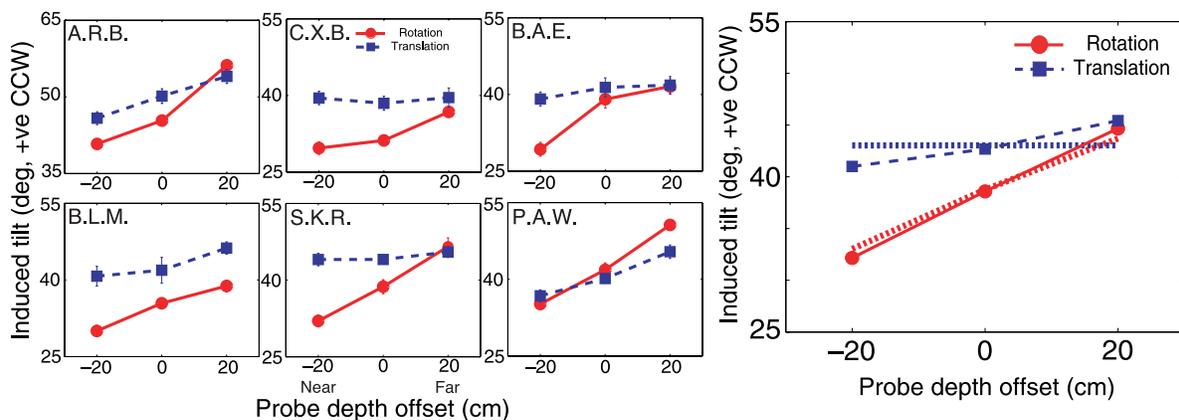


Figure 4. (Left) Results of Experiments 1 and 2 for six observers. Mean induced tilt is plotted as a function of probe depth. For rotations, induced tilt varies as a function of probe depth. For translations, induced tilt shows considerably less dependence on distance. (Right) Results of Experiments 1 and 2 for composite observer in rotation and translation conditions (mean of six observers in Experiments 1 and 2). A model of perceived induced tilt under the flow parsing hypothesis for the two conditions is also shown (dotted lines).

(dotted lines): Induced tilt was approximated using simple trigonometry and assuming vector summation of the horizontal (due to flow parsing) and vertical (due to physical movement) components of probe motion. The model is described in the following equation and has been fitted to the composite data using two parameters ( $G_R$  and  $G_T$ ), corresponding to different multiplicative gain factors for the horizontal component of perceived probe motion in the rotation and translation conditions, respectively (for derivation, see [Appendix A](#)).

$$\alpha_j = \tan^{-1}(G_j \theta_h / \theta_v); \quad j = R, T. \quad (3)$$

The best fitting gain factors (which minimized RMS error) were 0.53 (fits for individual observers range from 0.42 to 0.72) and 0.42 (range 0.36–0.54) for the rotation and translation conditions, respectively. To put these figures in context, the typical gain observed in experiments of judgment of locomotor heading direction from optic flow (perceived heading angle/actual heading angle) is around 0.5 (see Figure 2 of Lappe et al., 1999).

The results in [Figure 4](#) indicate that our observers' trajectory settings were consistent with the flow parsing hypothesis. This assertion was formally tested by conducting a mixed effects ANOVA (with subject treated as a random effect). Simulated movement type,  $M = \{\text{Rotation, Translation}\}$ , physical probe direction,  $p = \{V, V \pm 15^\circ, V \pm 30^\circ\}$  and probe depth,  $D = \{0.85 \text{ m}, 1.05 \text{ m}, 1.25 \text{ m}\}$  were treated as fixed effects. The analysis revealed a significant main effect of simulated movement type  $M$ ,  $F(1, 5) = 7.14$ ,  $p < .05$ , and depth  $D$ ,  $F(2, 10) = 29.28$ ,  $p < .001$ , but no effect of probe direction  $P$ ,  $F(4, 40) = 1.4$ , *ns*. Most importantly for the hypothesis proposed here, the analysis showed a highly significant two-way interaction ( $D \times M$ ) between depth and movement type,  $F(2, 40) = 20.01$ ,  $p < .001$ , indicating that there was strong evidence that as the probe depth varied observers responded differently in the two different observer movement conditions. Furthermore, there was no evidence for the three-way interaction  $S \times D \times M$  between the random effect subject,  $S$ , and the fixed effects  $D$  and  $M$ ,  $F(10, 1260) = 0.83$ , *ns*. This result suggests that the differences in patterns of the  $D \times M$  interaction observed for the six subjects (see [Figure 3](#)) are due to random variation and are not significantly different. No other treatment effects were found to be significant.

Subsequent tests on the treatment means indicated that consistent with our hypothesis, the  $D \times M$  interaction was driven by a large difference (around  $12.5^\circ$ ) in responses as a function of depth in the rotation condition and by a much smaller effect of depth on responses in the translation condition (around  $4.4^\circ$ ).

Explanatory theories of classical induced motion are numerous (for a review, see Reinhardt-Rutland, 1988). Of those, only local motion contrast could account for the

results reported here: If the motion of the probe could be compared to just those scene objects at a similar distance, then there would be relative motion between the probe and the scene. The relative motion would be compatible with the observers' responses. Due to our previous results, however, investigating flow parsing in 2D stimuli without local interactions between probe and background (Warren & Rushton, 2004), we do not believe that this account can explain the results presented here. In the following experiment we test this assertion.

## Experiment 3

A further experiment was designed to explore the nature of the mechanism driving the perception of object trajectory during self-movement. In particular, we were interested in whether the mechanism relied upon local interactions between target and background objects or a more global mechanism as might be expected if optic flow processing were implicated. Therefore, we were looking for evidence that the effect was not purely attributable to motion contrast between objects within a limited depth range.

As noted above readers familiar with the induced motion literature (e.g., see Post, Shupert, & Liebowitz, 1984; Reinhardt-Rutland, 1988) might be tempted to try to describe the results of [Experiments 1 and 2](#) with reference to previous findings. The results in the rotation condition could indeed be interpreted as the classic induced motion effect (Duncker, 1929). In the case of the observer translation condition, the results are perhaps slightly more difficult to explain by an induced motion account although some investigators have claimed that background motion in the fixation plane drives induced motion phenomena (e.g., Heckmann & Howard, 1991) and, to a certain extent, this would produce results consistent with those presented here.

Therefore, in a third experiment we investigate perceived trajectory when the probe is spatially separated from the background motion. No previous work on induced motion has examined such a complex situation, and to our knowledge no model of induced motion would be able to predict a pattern of results consistent with our parsing predictions and the experimental data. Specifically, we predict that if a global process such as flow parsing is implicated then, in spite of the spatial separation of probe and cubes, the results in this experiment will be similar to those found in [Experiments 1 and 2](#).

## Methods

The procedure and stimuli in this experiment were similar to those from [Experiments 1 and 2](#), except that the viewing distance was increased to 115 cm. Observers

were asked to report the perceived trajectory of a probe presented at different depths while viewing an array of cubes moving in a manner consistent with observer movement. However, in this experiment the probe was not always contained within the bounding volume of the cubes and could be isolated locally (see Figure 5A).

The depth range of the array of cubes was reduced to 20 cm, and the center of the volume was presented at 97, 115, or 133 cm from the observer which coincided with the three probe distances (a schematic representation is given in Figure 5A). Consequently, in the most extreme conditions the probe was spatially separated in depth from the nearest cube by at least 26 cm (i.e., with equivalent relative disparity between the two eyes of around 40 min arc). Once again observers saw stimuli that were consistent with simulated observer rotation or observer translation.

Because there is a suggestion of individual differences between observers and we are specifically attempting to compare two prospective mechanisms for optic flow processing, we sought to increase our likelihood of discriminating between them by imposing an inclusion criterion. We only tested observers (S.K.R. and B.A.E.) from Experiments 1 and 2 who showed a particularly strong influence of optic flow (and thus potentially a greater propensity toward flow parsing). In addition, two further observers, who did not take part in Experiments 1 and 2, were recruited. We adopted an inclusion criterion for these observers, requiring that when tested in the translation condition of Experiment 2, their trajectory settings demonstrated little or no dependence on probe depth (similar to B.A.E. and S.K.R.).

## Results and discussion

The important results from Experiment 3 are summarized in Figure 5B, which shows the average gradient (over observers) arising from linear fits to the induced tilt data in the rotation and translation conditions. If flow parsing were still occurring even in the presence of spatial separation of the probe and cubes, then similarly to Experiments 1 and 2 we would expect the gradient to be zero in the translation conditions but nonzero in the rotation conditions. Results are in line with this prediction for all but the “far” cube array condition with simulated observer translation (a potential explanation for this inconsistency is given below).

Induced tilts for both a representative observer (top row) and the composite observer obtained by averaging across participants (bottom row) are shown in Figure 5C. In the panels marked “near,” the volume was centered around the near probe distance so there were no cubes in the immediate vicinity when the probe was at the middle and far distances (condition marked “near” in Figure 5A). For the translation condition, the pattern of results shows little evidence for the dependence of induced tilt on distance. A similar pattern of results is found in the “middle” conditions of Figure 5C (condition marked “middle” in Figure 5A). Taken together, these data show that even when there is spatial separation between the cubes and the probe, there is still evidence for flow parsing. These data indicate that the results consistent with flow parsing are unlikely to be due to simple “motion contrast” (i.e., local differences in speed between the

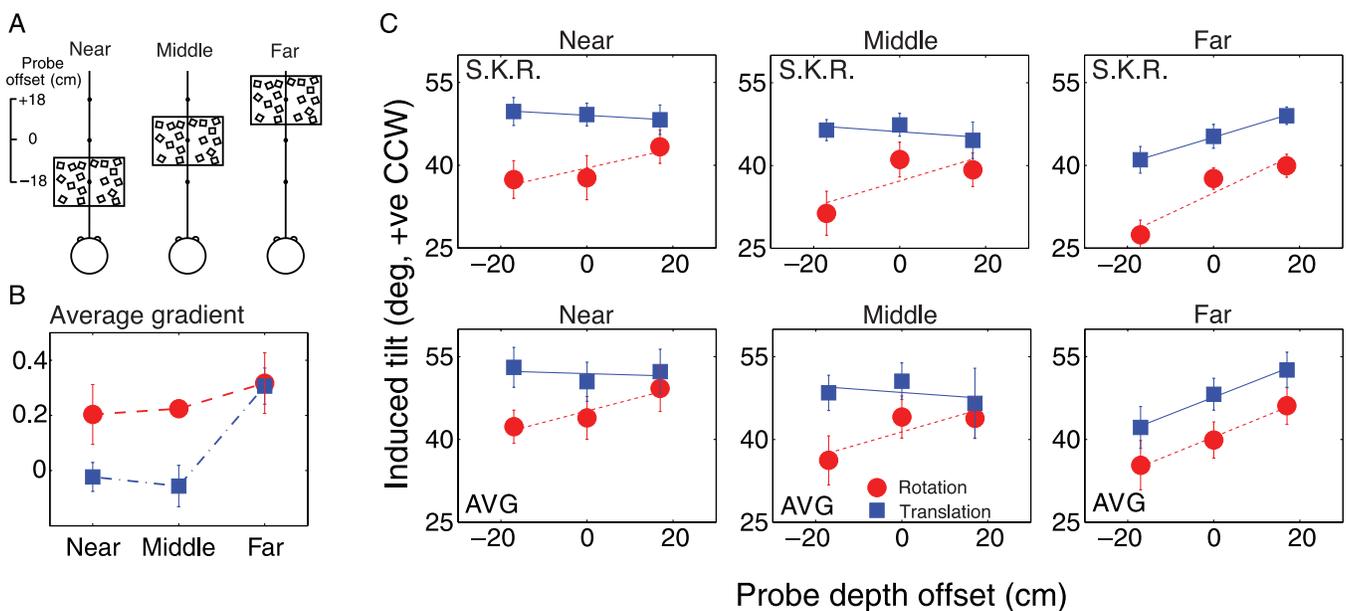


Figure 5. Experiment 3 set-up and results. (A) Schematic illustration of the conditions in Experiment 3. (B) Mean gradient of linear fit to rotation (circles) and translation (squares) data over all observers. Error bars show standard errors calculated over individual observer gradients. (C) Mean induced tilt as a function of probe depth when cube array was centered at near (−18 cm), middle (0 cm), and far (+18 cm) depths. Top row shows data for representative observer S.K.R., bottom row shows the data for the average observer (AVG).

probe and the elements immediately surrounding it) but rather involve a more global process.

In the “far” condition, the induced tilt shows a marked dependence on probe depth in both the rotation and translation conditions. This result seems at odds with the predictions of the flow parsing hypothesis. However, there is a plausible explanation for why this pattern of results was obtained. Variation in retinal speed, which distinguishes retinal motion due to lateral translation from retinal motion due to rotation, decreases with the distance of the volume; so too does the difference in binocular disparity that indicates the depth order of the cubes. Therefore, as the distance of the cubes increases, so does the potential for ambiguity about the type of self-movement that generated the retinal motion. In other words, in the “far” condition, we propose that the stimulus information was not sufficiently rich to resolve the ambiguity between self-movement types; the translation condition appeared very similar to the rotation condition. This is consistent with previous work which indicates that depth order information is important for flow parsing to occur (Rushton et al., 2006; Rushton & Warren, 2005; Warren & Rushton, 2004).

In spite of the spatial separation of the probe and cubes in this experiment, the results are still consistent with flow parsing in the “near” and “middle” conditions. We suggested that a local motion contrast account of induced motion could be compatible with the results of Experiments 1 and 2. However, the results of Experiment 3 suggest a mechanism that is global in nature. Therefore, we can rule out all classical induced motion explanations.

## Discussion

The results of the experiments reported here offer evidence that human observers use flow parsing to aid in the estimation of object trajectory during self-movement.

Participants showed individual differences in their responses, which varied in apparent consistency with the flow parsing hypothesis: Some observers showed clearer evidence for visual parsing of retinal flow, others showed less of a propensity toward flow parsing. Future experiments might examine whether the variability reflects broad differences in flow processing, which might be indicated by finding a correlation between flow parsing performance and performance on tasks such as judging heading from optic flow. Alternatively, the differences in observer responses might be due to other individual factors that interacted with the choice of stimulus parameters in this study. For example, in our stimulus, depth order was indicated by binocular disparity, and it has been reported that there is considerable variability in ability to perceive crossed and uncrossed disparities (e.g., van Ee & Richards, 2002).

In the experiments presented here, observers made judgments in the presence of optic flow patterns that signaled only one component of observer movement. In a separate study, we have extended Experiments 1 and 2 and found that these results are not limited to the simple flow fields described above. Target trajectories that were consistent with flow parsing were also found for more complex (and natural) flow fields resulting from a combination of observer rotation and translation (Warren & Rushton, 2004).

Note that the displays used here were cluttered scenes with objects at a range of distances defined by relative disparity. In previous experiments, we have examined the importance of depth information in the parsing process. When the depth order cue (binocular disparity in our experiments) is removed from such displays, the perceived target trajectory is no longer consistent with the flow parsing hypothesis (Warren & Rushton, 2004). Similarly, in the “far” condition of Experiment 3, the range of disparities is restricted and the flow parsing appears to break down. These results may explain why other investigators who have addressed similar questions have reached different conclusions from those presented here (Brenner, 1991; Harris, Jenkin, Dyde, & Jenkin, 2004).

Given the short display period we used (2 s), we did not expect observers to experience vection (for an indication of typical vection onset times with a range of stimuli, see Palmisano, Burke, & Allison, 2003) and none reported doing so. Therefore, one interesting conclusion we can draw is that whatever changes occur during vection, they do not appear to be necessary for flow parsing.

Reinhardt-Rutland (1988) provides a useful review of the explanatory theories of induced motion. He groups the theories into the following categories: Duncker’s (1929/1938) Theory; Alteration of the Observer’s Perception of Space; Felt and Cancelled Eye Movements; Induced Movement and “Intelligent” Perception; and Sensory and Neural Processes. None of the first four explanations are consistent with the results we report because changing the probe depth should not change the perceived horizontal probe motion (and hence induced tilt) under these accounts. The final account (Sensory and Neural Processes) indicates that the illusory motion is seen as a consequence of local motion contrast arising from lateral inhibition in motion detectors. It is possible that this account could explain the results found in Experiments 1 and 2. However, this account would not predict or explain the results of Experiment 3. We suggest, instead, that the results of all three experiments can be explained by a single parsimonious theory, namely, the flow parsing hypothesis.

Our third experiment was motivated in part by an attempt to identify where flow parsing is performed. On the basis of animal studies, MST would appear the most likely candidate. However, recent imaging work (Smith, Wall, Williams, & Singh, 2005) has cast some doubt on whether the animal neurophysiology provides a good model for the organization of higher levels of human motion processing. Therefore, in

this paper we restrict ourselves to a computational description. The results of the third experiment point toward a contribution of global motion processing; the perceived motion of the probe was influenced by nonlocal motion, that is, motion in other depth planes.

A related problem to that studied here is the perception of heading when objects are moving within the scene. If the observer requires an accurate estimate of heading then retinal motion due to object and self-movement must be separated out. Royden (2002) proposed detectors that could perform this task, and others have also worked on this problem (Hildreth, 1992). Further models have been suggested that aim to detect object movement during observer movement (Thompson & Pong, 1992). These studies may provide an important starting point for modeling the flow parsing process proposed here.

Most previous work on optic flow has been motivated by the hypothesis that the brain identifies patterns of retinal motion for guiding locomotion through the environment (Gibson, 1950; Warren & Hannon, 1988 but see Rushton, Harris, Lloyd, & Wann, 1998). We interpret our findings in the context of a different body of work on optic flow. Miles and colleagues have suggested a role for flow processing in driving short-latency, reflexive eye movements to minimize retinal movement and thus stabilize objects of interest on the retinae during self-movement (e.g., Bussetini et al., 1997; Miles, 1997). The results of this study and other recent work (Rushton et al., 2006; Rushton & Warren, 2005) suggest a further complementary use for flow processing. After the oculomotor systems have stabilized the retinal images as much as possible through eye movements, flow parsing then separates the remaining retinal motion into components due to self-movement and object movement. This work supports an emerging role for optic flow in stabilization and updating of the visual world during self-movement.

## Appendix A

In this section, we describe the flow parsing model used to generate the predictions shown in Figure 4. Because the induced tilt measure is corrected for differences in physical probe trajectory between trials, this model is constructed as if the probe always moved in a purely vertical direction.

Let  $\theta_v$  be the retinal extent of the physical probe motion in the vertical direction over the course of a single 2-s trial. As noted in the methods, the probe moves at a speed of 0.9 cm/s. This corresponds to a retinal movement of 0.83°, 0.98°, and 1.21° at the far, middle, and near probe distances, respectively.

If the visual system were able to undertake perfect flow parsing in our experiments, then all of the background retinal motion should be attributed to self-movement. As a consequence, this component would be parsed from the

retinal image and the probe would also be seen to move horizontally by an amount  $\theta_h$ , which depends upon the type of movement undertaken and the extent of this movement. The observer movements simulated in this experiment correspond to a head rotation of 0.75 deg/s and a head translation of 2 cm/s. Under perfect flow parsing, these movements should then generate a perceived  $\theta_h$  of 1.5° in the rotation condition (irrespective of probe distance) and 1.83°, 2.18°, and 2.69° in the translation condition for the far, middle, and near probe distances, respectively.

Using simple trigonometry and assuming vector summation of the horizontal (due to flow parsing) and vertical (due to physical movement) components of probe motion, the induced tilt can then be predicted as

$$\alpha = \tan^{-1}(G\theta_h/\theta_v). \quad (\text{A1})$$

Here  $G$  is a gain parameter reflecting the fact that observers do not parse flow perfectly in our experiment. This may be due to a propensity to use extra-retinal information or simply the fact that our stimuli were not sufficiently rich to engage the flow parsing mechanism.

The predictions presented in Figure 4 were generated using this model with the values for  $\theta_h$  and  $\theta_v$  described above and different gain components for the rotation and the translation conditions that were fitted to the data. If the flow parsing mechanism is underpinned by neural regions sensitive to flow then because there are thought to be neural regions that process different types of flow components (e.g., rotation/translation), it is possible that the subsystems have developed separate gain components.

## Acknowledgments

We thank Tom Freeman, Mike Landy, and an anonymous referee for comments on an earlier draft of this paper. In addition, we thank Andrew Belyavin for insightful suggestions for the statistical analysis of Experiments 1 and 2.

Commercial relationships: none.

Corresponding author: Paul A. Warren.

E-mail: warrenpa@cardiff.ac.uk.

Address: School of Psychology, Tower Building, Cardiff University, Cardiff, Wales, UK.

## Footnote

<sup>1</sup> It was not possible to record eye movements through the shutter glasses; however, we have no reason to believe that observers chose to make the task more difficult for

themselves by looking at objects other than the target. Furthermore, moving the eyes would have generated additional flow components in the image that should also be parsed out under the flow parsing hypothesis. Perhaps most importantly, if observers chose to look around the display, it is difficult to imagine a pattern of behavior that might have led to the systematic results that were obtained.

## References

- Brecher, G. A., Brecher, M. H., Kommerell, G., Sauter, F. A., & Sellerbeck, J. (1972). Relation of optical and labyrinthian orientation. *Optica Acta*, *19*, 467–471.
- Brenner, E. (1991). Judging object motion during smooth pursuit eye movements: The role of optic flow. *Vision Research*, *11*, 1893–1902. [[PubMed](#)]
- Bussettini, C., Masson, G. S., & Miles, F. A. (1997). Radial optic flow induces vergence eye movements with ultra-short latencies. *Nature*, *390*, 512–515. [[PubMed](#)]
- Duncker, K. (1929). Über induzierte Bewegung. *Psychologische Forschung*, *12*, 180–259.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Gogel, W. C. (1990). A theory of phenomenal geometry and its applications. *Perception & Psychophysics*, *48*, 105–123. [[PubMed](#)]
- Harris, L. R., Jenkin, M. R., Dyde, R. T., & Jenkin, H. L. (2004). Failure to update spatial location correctly using visual cues alone [[Abstract](#)]. *Journal of Vision*, *4*(8):381, 381a, <http://journalofvision.org/4/8/381/>, doi:10.1167/4.8.381.
- Heckmann, T., & Howard, I. P. (1991). Induced motion: Isolation and dissociation of egocentric and vection-entrained components. *Perception*, *20*, 285–305. [[PubMed](#)]
- Hildreth, E. C. (1992). Recovering heading for visually guided navigation in the presence of self-moving objects. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *337*, 305–313.
- Howard, I. P., & Childerson, L. (1994). The contribution of motion, the visual frame and visual polarity to sensations of body tilt. *Perception*, *23*, 753–762. [[PubMed](#)]
- Johansson, G. (1974). Vector analysis in visual perception of rolling motion. A quantitative approach. *Psychologische Forschung*, *36*, 311–319. [[PubMed](#)]
- Lappe, M., Bremmer, F., & van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, *3*, 328–336. [[PubMed](#)]
- Lee, D. N., & Aronson, E. (1974). Visual proprioceptive control of standing in human infants. *Perception & Psychophysics*, *15*, 529–532.
- Lepecq, J. C., Jouen, F., & Dubon, D. (1993). The effect of linear vection on manual aiming at memorized directions of stationary targets. *Perception*, *22*, 49–60. [[PubMed](#)]
- Miles, F. A. (1997). Visual stabilization of the eyes in primates. *Current Opinion in Neurobiology*, *7*, 867–871. [[PubMed](#)]
- Palmisano, S., Burke, D., & Allison, R. S. (2003). Coherent perspective jitter induces visual illusions of self-motion. *Perception*, *32*, 97–110. [[PubMed](#)]
- Post, R. B., Shupert, C. L., & Liebowitz, H. W. (1984). Implications of OKN suppression for smooth pursuit for induced motion. *Perception & Psychophysics*, *36*, 493–498. [[PubMed](#)]
- Reinhardt-Rutland, A. H. (1988). Induced movement in the visual modality: An overview. *Psychological Bulletin*, *103*, 57–71. [[PubMed](#)]
- Royden, C. S. (2002). Computing heading in the presence of moving objects: A model that uses motion-opponent operators. *Vision Research*, *42*, 3034–3058. [[PubMed](#)]
- Royden, C. S., Wolfe, J. M., & Klempen, N. (2001). Visual search asymmetries in motion and optic flow fields. *Perception & Psychophysics*, *63*, 436–444. [[PubMed](#)] [[Article](#)]
- Rushton, S. K., Bradshaw, M. K., & Warren, P. A. (2006). The pop out of scene-relative object movement against retinal motion due to self movement. *Cognition*, *105*, 237–245. [[PubMed](#)]
- Rushton, S. K., Harris, J. M., Lloyd, M. R., & Wann, J. P. (1998). Guidance of locomotion on foot uses perceived target location rather than optic flow. *Current Biology*, *8*, 1191–1194. [[PubMed](#)] [[Article](#)]
- Rushton, S. K., & Warren, P. A. (2005). Moving observers, relative retinal motion and the detection of object movement. *Current Biology*, *15*, R542–R543. [[PubMed](#)] [[Article](#)]
- Smith, A. T., Wall, M. B., Williams, A. L., & Singh, K. D. (2005). Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience*, *23*, 561–569. [[PubMed](#)]
- Swanston, M. T., & Wade, N. J. (1988). The perception of visual motion during movements of the eyes and of the head. *Perception & Psychophysics*, *43*, 559–566. [[PubMed](#)]
- Swanston, M. T., Wade, N. J., Ono, H., & Shibuta, K. (1992). The interaction of perceived distance with the perceived direction of visual motion during movements of the eyes and the head. *Perception & Psychophysics*, *52*, 705–713. [[PubMed](#)]

- Tcheang, L., Gilson, S. J., & Glennerster, A. (2005). Systematic distortions of perceptual stability investigated using immersive virtual reality. *Vision Research*, *45*, 2177–2189. [[PubMed](#)]
- Thompson, W. B., & Pong, T. (1992). Detecting moving objects. *Computer Vision*, *4*, 39–57.
- van Ee, R., & Richards, W. (2002). A planar and a volumetric test for stereoaomaly. *Perception*, *31*, 51–46. [[PubMed](#)]
- Van Pelt, S., & Medendorp, W. P. (2007). Gaze-centered updating of remembered visual space during active whole-body translations. *Journal of Neurophysiology*, *97*, 1209–1220. [[PubMed](#)]
- Wallach, H. (1987). Perceiving a stable environment when one moves. *Annual Review of Psychology*, *38*, 1–27. [[PubMed](#)]
- Warren, P. A., & Rushton, S. K. (2004). Optic flow and the induced motion illusion [[Abstract](#)]. *Journal of Vision*, *4*(8):893, 893a, <http://journalofvision.org/4/8/893/>, doi:10.1167/4.8.893.
- Warren, W. H., & Hannon, D. J. (1988). Direction of self motion is perceived from optical-flow. *Nature*, *336*, 162–163.
- Zhou, H. H., Wei, M., & Angelaki, D. E. (2002). Motor scaling by viewing distance of early visual motion signals during smooth pursuit. *Journal of Neurophysiology*, *88*, 2880–2885. [[PubMed](#)] [[Article](#)]