Laurence R. Harris · Michael Jenkin
Daniel C. Zikovitz

# Visual and non-visual cues in the perception of linear self motion

**Abstract** Surprisingly little is known of the perceptual consequences of visual or vestibular stimulation in updating our perceived position in space as we move around. We assessed the roles of visual and vestibular cues in determining the perceived distance of passive, linear self motion. Subjects were given cues to constant-acceleration motion: either optic flow presented in a virtual reality display, physical motion in the dark or combinations of visual and physical motions. Subjects indicated when they perceived they had traversed a distance that had been previously given to them either visually or physically. The perceived distance of motion evoked by optic flow was accurate relative to a previously presented visual target but was perceptually equivalent to about half the physical motion. The perceived distance of physical motion in the dark was accurate relative to a previously presented physical motion but was perceptually equivalent to a much longer visually presented distance. The perceived distance of self motion when both visual and physical cues were present was more closely perceptually equivalent to the *physical* motion experienced rather than the simultaneous visual motion, even when the target was presented visually. We discuss this dominance of the physical cues in determining the perceived distance of self motion in terms of capture by non-visual cues. These findings are related to emerging studies that show the importance of vestibular input to neural mechanisms that process self motion.

L.R. Harris (✉)
Department of Psychology, York University, Toronto, Ontario, M3J 1P3, Canada
e-mail: harris@yorku.ca
Tel.: +1-416-7362100, Fax: +1-416-7365814

M. Jenkin
Department of Computer Science, York University, Toronto, Canada

L.R. Harris · M. Jenkin · D.C. Zikovitz
Centre for Vision Research at York University and
Department of Biology, York University, Toronto, Canada

## Introduction

As we move around the world we have a seemingly effortless perception of the extent of our movement. As we walk along a corridor to a door, for example, we stop at the right place and do not overshoot or undershoot our target. How do we manage this task? There are many cues that could contribute to the perception of how far we have moved, including visual, vestibular, other proprioceptor information, such as the movements of the limbs, and re-afference. This paper investigates the relative contributions of optic flow and passive, non-visual physical cues.

Much attention has been paid in recent years to optic flow as a visual cue to motion (Gibson 1950; Koenderink and van Doorn 1975; Royden et al. 1992; see Lappe 2000). When moving through a three-dimensional environment, the components of the retinal image stream across the retina and the resulting optic flow contains information about the direction and, in the presence of scaling information, the velocity and magnitude of the movement (Gibson 1950). It has been demonstrated that people can use optic flow to assess their direction of travel (Warren et al. 1988, 1991; Royden et al. 1992; Lappe and Rauschecker 1994). But, although it has been shown that honey bees can use optic flow to assess distance of travel (Esch and Burns 1995, 1996; Srinivasan et al. 1997, 2000) and that humans can discriminate differences in distances based exclusively on optic flow (Bremmer and Lappe 1999), there have been no explicit claims or demonstrations that humans can estimate their distance of travel from optic flow.

Another important cue to linear movement is provided by the otoliths of the vestibular system, probably supplemented by somatic graviceptors (Mittelstaedt 1997). The otolith system transduces only linear acceleration (Lowenstein 1974), so periods of constant velocity cannot be registered by this system, but, with this caveat, position rela-

tive to some initial state can theoretically be obtained by two integrations over time. People can use vestibular information to assess a position change (Mayne 1974; Parker et al. 1979; Israël et al. 1993; Loomis et al. 1993; Glasauer et al. 1994; Berthoz et al. 1995) and their direction of travel (Telford et al. 1995; Ohmi 1996). But humans' ability to assess their distance of travel from non-visual cues in the presence or absence of optic flow has been a surprisingly neglected area of investigation, perhaps partly because of the technical difficulties in dissociating these cues before the advent of virtual reality technologies.

Areas of the brain that are involved in representing space have inputs from both the visual and vestibular systems (parietal: Andersen et al. 1997, hippocampus: Sharp et al. 1995; Smith 1997, visual cortex: Vanni-Mercier and Magnin 1982) but surprisingly little is known quantitatively about the perceptual consequences of either visual or vestibular stimulation as we move around. We therefore measured the distance of the perceived self motion resulting from visual or vestibular stimulation alone or in combination. These experiments provide an important reference for interpreting emerging studies of multimodal convergence in the parietal and hippocampal areas (Duffy 1998; Bremmer et al. 1999). Some of the experiments reported here have appeared in abstract form (Harris and Jenkin 1996; Zikovitz et al. 1998; Redlick et al. 1999).

## Materials and methods

### Choice of method

Measuring how far someone perceives themselves to have moved presents some interesting methodological considerations. Simply asking people to estimate how far they have moved requires them to make a relative judgement against an internal representation of some kind of yardstick. Distortions in the representation of the yardstick, such as stimulus compression or expansion (Stevens 1955; Parker et al. 1979) when judging multiples of the yardstick, complicate the interpretation of such data. Such a technique cannot be used to predict the accuracy with which people perceive their movement through a particular given target distance. Asking subjects to reproduce previously travelled distances (Berthoz et al. 1995) also does not address the veridicality of perception since an inaccuracy or systematic bias in the perception of the initial distance may be matched by similar inaccuracies and bias in the measurement trials.

In this study we asked subjects to judge their motion relative to target distances presented through one of two modalities: either visually or by physical motion. For example, subjects were shown a target distance visually and asked to match this with physical motion. This cross-modal matching task cannot be performed by simply reproducing an experience but requires an internal representation of the stimulus. Trials in which bimodal stimuli (vision and physical motion) were matched to either visually or physically presented targets allow us to access the relative contributions of the sensory modalities without needing access to absolute-distance judgements.

### Procedure

For each trial, subjects were first presented with a reference distance either visually or by physical demonstration (described below). They were then exposed to passive, constant-acceleration movement which consisted of either: (a) physical movement in the dark, (b) visual motion only or (c) a combination of visual and physical motion. The sequence of trial types was selected randomly but the subjects' task was always the same: to indicate by pressing a button when they perceived themselves to have travelled through the reference distance. Subjects were not provided with feedback concerning their performance at any time. At the point they pressed the button, subjects were indicating that they *perceived* they had travelled through the target distance. The *actual* motion they had experienced at this point was then taken as perceptually equivalent to this perceived distance.

Experiments were approved by the York University Ethics Approval Committee. Subjects were paid for their participation at standard York subject rates. The number of subjects used for each part of the study is indicated in Table 1.

**Table 1** Summary of the perceptual gains (perceived/actual motion) measured under the various conditions of this study. The conditions for section 2 correspond to different accelerations used for the target (in which the distance was presented by physical demonstration) and the subsequent trial. The conditions for sections 5 and 6 correspond to the ratio of visual to physical motion. For sections 5 and 6 there are two perceptual gains corresponding to whether the perceived motion is expressed as a fraction of the actual visual motion or the actual physical motion

| Results section | Target | Trial | Condition | Number of subjects | Figure | Perceptual gain (perceived motion/actual motion) | | Regression coefficient |
|---|---|---|---|---|---|---|---|---|
| 1 | Visual | Physical | Real targets | 12 | 3 | 2.13 | | 0.79 |
| | Visual | Physical | Virtual targets | 10 | 3 | 2.00 | | 0.67 |
| 2 | Physical | Physical | Fast/slow | 7 | 3 | 0.88 | | 0.75 |
| | Physical | Physical | Slow/fast | 7 | 3 | 1.11 | | 0.89 |
| | Physical | Physical | Slow/slow | 7 | 3 | 1.17 | | 0.85 |
| | Physical | Physical | Fast/fast | 7 | 3 | 0.99 | | 0.87 |
| 3 | Visual | Visual | Vision only | 9 | 4 | 0.96 | | 0.87 |
| 4 | Physical | Visual | Vestibular target | 3 | 4 | 0.23 | | 0.75 |
| | | | | | | (Perceived) motion/visual motion | (Perceived motion/physical motion) | |
| 5 | Visual | Visual and physical | Vis/vest = 0.5 | 12 | 5 | 7.15 | 3.45 | 0.84 |
| | Visual | Visual and physical | Vis/vest = 1.0 | 12 | 5 | 3.57 | 3.57 | 0.88 |
| | Visual | Visual and physical | Vis/vest = 2.0 | 12 | 5 | 2.70 | 5.56 | 0.93 |
| 6 | Physical | Visual and physical | Vis/vest = 0.5 | 4 | 6 | 1.64 | 0.83 | 0.49 |
| | Physical | Visual and physical | Vis/vest = 1.0 | 4 | 6 | 2.33 | 2.33 | 0.27 |
| | Physical | Visual and physical | Vis/vest = 2.0 | 4 | 6 | 1.00 | 2.0 | 0.37 |

# PRESENTING "VISUAL" TARGETS



**A** *Subject shown target distance in helmet* **B** *"Push button when you reach target"* **C** *Subject's eye view of the virtual corridor.*
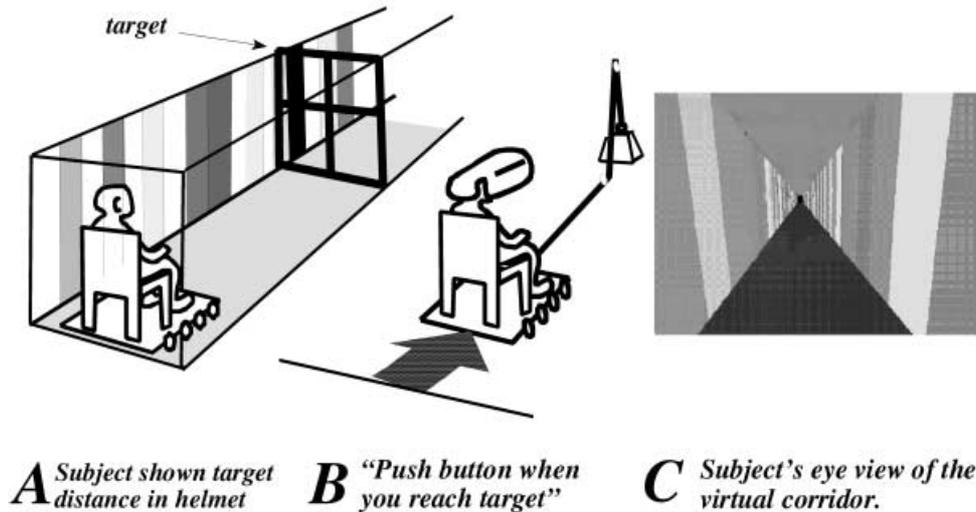
**Fig. 1A–C** Experimental set up. Subjects wore a virtual reality helmet and sat on a cart which was attached by a rope to a weight hung from pulleys. When the weight was released it pulled the cart at a constant acceleration. The acceleration could be varied by varying the size of the weight. The helmet displayed a virtual environment in which subjects saw a striped, virtual corridor with grey ceiling and floor as illustrated in **C**. To present targets visually, a cross-frame was displayed at some distance down the corridor (**A**). Subjects were encouraged to move their heads to obtain parallax and perspective cues to help them assess the target distance. Following target presentation, trials could be either physical motion in the dark (**B**), visual motion only (in which the corridor shown in **C** was moved past the subject, without the target) or a combination of the two

## Physical motion equipment

Subjects were subjected to physical motion by sitting on a chair that was mounted on a mobile cart (see Figs. 1, 2). The cart rolled on low-friction, in-line skate wheels and ran on a smooth floor. The cart was attached by a rope through an arrangement of pulleys to a weight which could be dropped either from a frame (1.5 m) or down a stairwell (8 m). The distance of the drop on the frame was converted to the equivalent of a drop of up to 4.5 m by the pulley arrangement. The force of the dropping weight propelled the cart across the floor at a constant acceleration of between 0.08 and 0.54 m.s$^{-2}$ depending on the mass of the weight and the subject.

For the virtual-visual display conditions, the cart position was transduced by running a thin, earth-fixed wire around the optical-encoder shaft of a mouse mounted on the cart. The system was calibrated by moving the cart through known distances by hand to obtain the calibration factor between rotations of the mouse shaft and metres travelled (resolution: 1/30 cm). The calibrated signal from the mouse was sampled and stored by an SGI (Silicon Graphics) O2 computer which also generated the visual display (see below) and recorded subject responses.

Experiments not involving a virtual reality display were carried out in a corridor with the movement of the cart powered by weights that were arranged to fall down the centre of a three-storey stairwell in the Chemistry and Computer Science Building of York University. The experiments allowed greater flexibility in distance (up to 8 m) and higher rates of acceleration (up to 0.54 m.s$^{-2}$). For these experiments distances were measured by a tape measure attached to the cart by a quick-release clip and time of travel by subject- and experimenter-controlled stopwatches. When subjects felt they had travelled through the appropriate dis-

## PRESENTING "VESTIBULAR" TARGETS



**A** *Subject dragged through target distance*
**B** *Subject returned to start* **C** *"Push button when you reach target"*
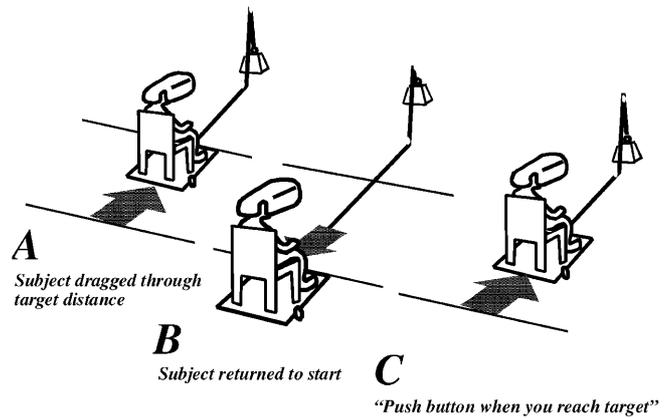
**Fig. 2A–C** To present target distances physically, subjects were moved through the target distance in total darkness at constant acceleration by dropping a weight attached to the cart (**A**). They were stopped suddenly at the target distance and returned to the start position (**B**). Following target presentation, trials commenced (**C**) as described in Fig. 1

tance they stopped their stopwatch and vocalised. The experimenter's watch was then stopped and the tape jammed, thus activating the quick release from the cart. Reaction times were estimated at less than 100 ms during which time the error in distance, even at the longest distances (8 m) and highest accelerations (0.54 m.s$^{-2}$) was less than 4%.

## Visual display equipment

During visual presentations, subjects viewed a 84°×65° display presented on a single-screen, non-stereoscopic head-mounted display (Liquid Image MRG3). The display simulated a virtual corridor 50 m long, 2 m wide and 2.5 m high, whose dimensions were based on the dimensions of a typical corridor at York University (Fig. 1). The walls of the corridor were painted with multicoloured, vertical stripes 0.5 m wide which changed colour on a random schedule. The changing colour reduced the possibility that subjects tracked the stripes. The floor and ceiling were black. The image was displayed at optical infinity. The measured position and

orientation of the subject's head was used to generate the appropriate view in the display such that, as the subjects moved their heads, they saw the view as if they were in a real corridor. Thus they had parallax and perspective cues concerning the dimensions of the simulated corridor.

The position and orientation of the subject's head was measured by a six degree-of-freedom Flock of Birds head tracker (Ascension Technologies) attached to the helmet. The Flock of Birds reported the orientation and linear position of the receiver (each with three degrees of freedom) relative to a reference transmitter (resolution $0.5°$ and 0.18 cm) which in this case was mounted on the cart. During physical motion, the position of the subject's head relative to the world was derived from the vector sum of the subject's head position relative to the cart, measured by the Flock of Birds relative to its reference transmitter mounted on the cart, and the physical position of the cart relative to the room, measured by the earth-fixed wire (see above). This combined measurement technique was necessary because of the limited operational range of the Flock of Birds system (1.2 m). The subject's head was unrestrained and the subject was encouraged to look around the simulated corridor. An advantage of using the head-mounted visual display is that the visual displacement presented using virtual reality could be systematically decoupled from the actual displacement of the cart.

Movement in the virtual corridor for vision-only trials was created by recording the physical motion of the cart on a previous occasion and simulating it later to drive the virtual reality display. This ensured that the visual conditions were the same in the vision-only and the vision-plus-vestibular conditions (cf. Harris et al. 1981).

### Calibration of the visual display

It was important that perceived distances and scale in the visual display were correctly calibrated to the real world. To obtain accurate calibration of the head-mounted display's optics, we used an empirical method. Subjects were presented with a target at a simulated distance (for example, 2 m) and then lifted the helmet and viewed a real-world target of the same dimensions at the same distance. Subjects raised and lowered the helmet while the simulated focal length of the virtual reality display was adjusted until the simulated and real targets appeared to be at the same distance. Subjects were encouraged to move their heads around during this exercise to generate parallax cues. The match was verified at several distances. An additional confirmation of the effectiveness of the calibration came from the observation that there was no difference between using real-world and virtual visual stimuli as targets (see below).

### Presentation of visual reference distance

Visual reference distances were presented using either real-world targets or a head-mounted display system. Subjects were shown real targets at between 1.5 and 6 m and simulated targets at between 0.5 and 10 m down the virtual corridor in the head-mounted display (see Fig. 1 and insets to Fig. 3). For the real target, an assistant indicated the distance with their arm or a metre rule. The simulated target was a red frame that went all round the edges of the corridor with vertical and horizontal bars forming a cross. For both types of targets, subjects were encouraged to move their heads around to help get a good idea of how far away the target was, using parallax and perspective cues before subject motion commenced.

### Presentation of physical reference distance

To present a physical reference distance, subjects were moved at a constant acceleration in complete darkness through the target distance. The weight attached to the rope that pulled the cart was dropped, dragging the cart at constant acceleration through the

specified distance (0.5–6 m). When the target distance was reached, an auditory cue was given and the cart was stopped suddenly by being grabbed by the experimenter. The cart was then moved back to the start point before the start of the experimental motion. The return was necessitated by our restricted track length and resulted in subjects actually being exposed to the reference distance twice (once on the way out and once on the way back). For experiments in which target distances were presented by physical movement that subjects then had to match by subsequent physical motion, the accelerations could be varied between two parts of the trial by adjusting the weight that pulled the cart.

## Results

### Physical motion in the dark perceptually equivalent to a visually presented target distance

When subjects moved at a constant acceleration in complete darkness they consistently and dramatically overestimated how far they had travelled relative to a previously presented visual target for both real and simulated targets (see Fig. 3). Since subjects pressed the button when they perceived they had travelled through the target distance, the initially presented 'target distance' corresponds to their perception. The perceived distance was plotted against the actual motion that they needed to achieve this perception. This is shown by the white and purple circles in Fig. 3 for real and virtual targets, respectively. The slope of such plots is defined as the **perceptual gain** (perceived/actual motion). The perceptual gain between visually presented target distance and the physical motion needed to achieve this perception was 2.13 for real targets and 2.0 for virtual targets, that is subjects pressed the button when they had in fact travelled through only about half the actual distance. A regression analysis showed that the difference between slope for judgments made using real targets and slope for judgments using virtual targets was not significant [$t$ (62)=1.67, n.s.]. The accelerations used here were from 0.1 to 0.3 m.s$^{-2}$. The data are summarised in Table 1.

### Physical motion in the dark perceptually equivalent to a physically presented target distance

When target distances were presented by physical demonstration (see Fig. 2), subjects were subsequently able to indicate fairly accurately when they had travelled in complete darkness through that distance (Fig. 3 filled symbols). In order to ensure that subjects were indeed matching distance estimates rather than merely waiting for the same period of time that the demonstration took, we combined different accelerations (and hence different times) for the demonstration and the matching portions of the trials. Trials consisted of either slow or fast physically presented targets, followed by either slow or fast matching movement. The four resulting combinations were randomly interleaved. Each condition was presented to each of seven subjects once. Slow accelerations
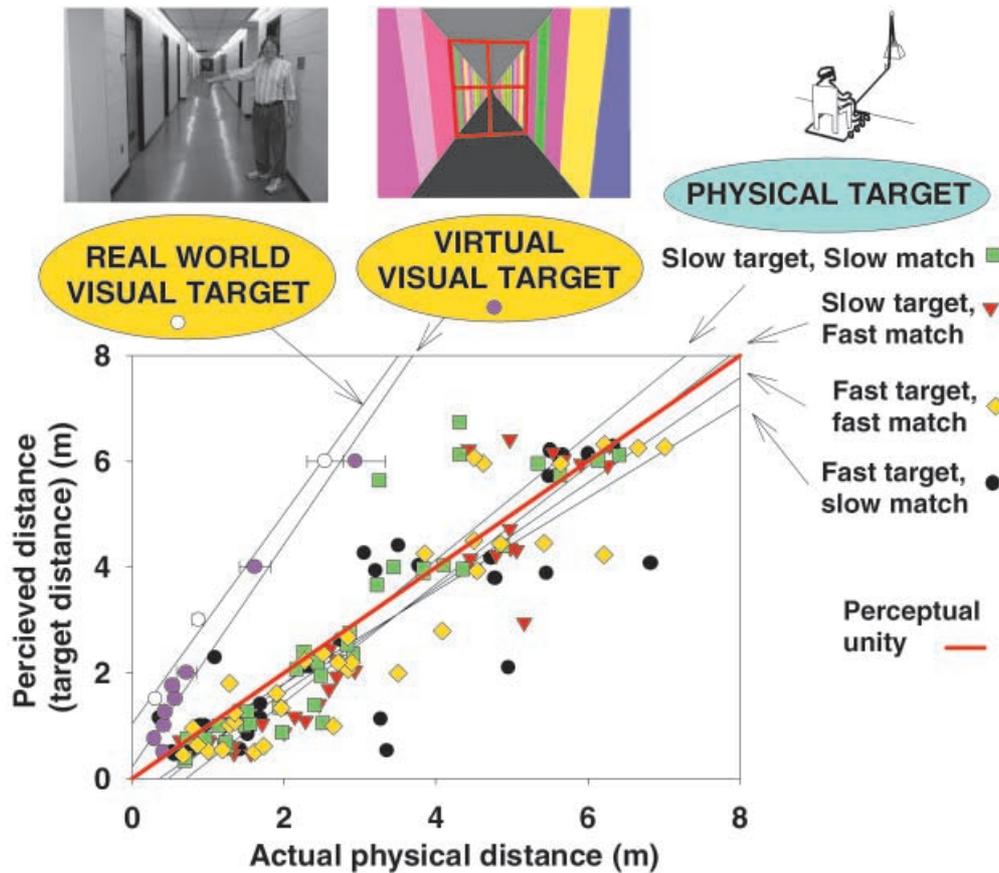
**Fig. 3** Judgements of the distance of self motion in the dark when physical motion was matched either to previously presented physical targets (labelled "physical target") or to visual targets (labelled "visual target"). The *horizontal axis* ("actual physical distance") is the actual physical distance the subjects had travelled in the dark at the point they pressed the button indicating that they perceived that they had travelled through the target distance ("perceived distance", *vertical axis*). The slope of such graphs describes the subjects' perceptual gain (perceived/actual motion). The *red line* indicates accurate performance: a perceptual gain of unity. Visual targets were either real (*open circles*) indicated by a person standing in a real corridor (*inset*) or virtual (*purple circles*) indicated by a cross-frame in a virtual corridor (*inset*). Physical targets (see Fig. 2) could be either fast (0.53 m.s⁻²) or slow (0.34 m.s⁻²) followed by either fast or slow motion (see *key*). *Error bars* for the visual-target data indicate standard errors. Data are the average of 10 subjects for the virtual visual targets and 12 subjects for the real world visual targets. Individual data points from 7 subjects are shown for the physically presented target responses

averaged 0.34 m.s⁻² and fast accelerations 0.53 m.s⁻². The accelerations used for each subject varied a little from these averages, but the ratio between fast and slow conditions was constant. Perceived distances (target distance) are plotted as a function of actual distance in Fig. 3 (green, red, yellow and black symbols).

When acceleration was higher in the demonstration part than in the subsequent matching part of the trial, the time taken to cover the target distance was correspondingly shorter in the demonstration part. Thus, if subjects were relying on a time estimate, they would press too early corresponding to a shorter actual dis-

tance, leading to steeper slopes and higher perceptual gains (1.54 using the average acceleration values). In contrast, when the acceleration was slower in the first part, times would be proportionally longer leading to later button presses and a shallower slope (0.65 using the average acceleration values). In fact, an ANOVA revealed no significant difference between any of the four slopes (slow/slow, fast/fast, slow/fast, fast/slow) [$F(3,126)=1.22$]. The perceptual gains varied from 0.88 to 1.17 with an average of 1.04. The data are summarised in Table 1.

## Vection perceptually equivalent to a visually presented target distance

When subjects were shown a visual target and then subjected to purely visual motion in the virtual reality display while sitting stationary on the cart they experienced a sense of vection in which they felt they were moving smoothly down the virtual corridor. These trials were randomly intermixed with trials in which they did experience physical motion. For the range of constant accelerations used in this study (0.1–0.5 ms⁻² subjects were accurate at judging the distance at which they reached the position of the previously presented visual target. The regression line of perceived against actual distance (Fig. 4 open circles) indicates a perceptual gain (perceived/actual motion) of 0.96 ($r^2$ 0.87).

**Fig. 4** Judgements of the distance of visually induced vection when the motion was matched either to previously presented physical targets (*closed circles*, *line* labelled "vestibular target") or to visual targets (*open circles*, *line* labelled "visual target"). Conventions and axes as for Fig. 3. The *dashed line* indicates a perceptual gain of unity. Although visually demonstrated targets could be accurately matched by subsequent visual motion at these constant accelerations (perceptual gain=0.96), physically presented targets had a very low perceptual gain (perceived/actual motion=0.23) indicating that much more visual motion was required to match them. Data are the averages and standard errors for nine subjects (visual targets) and three subjects (vestibular targets)
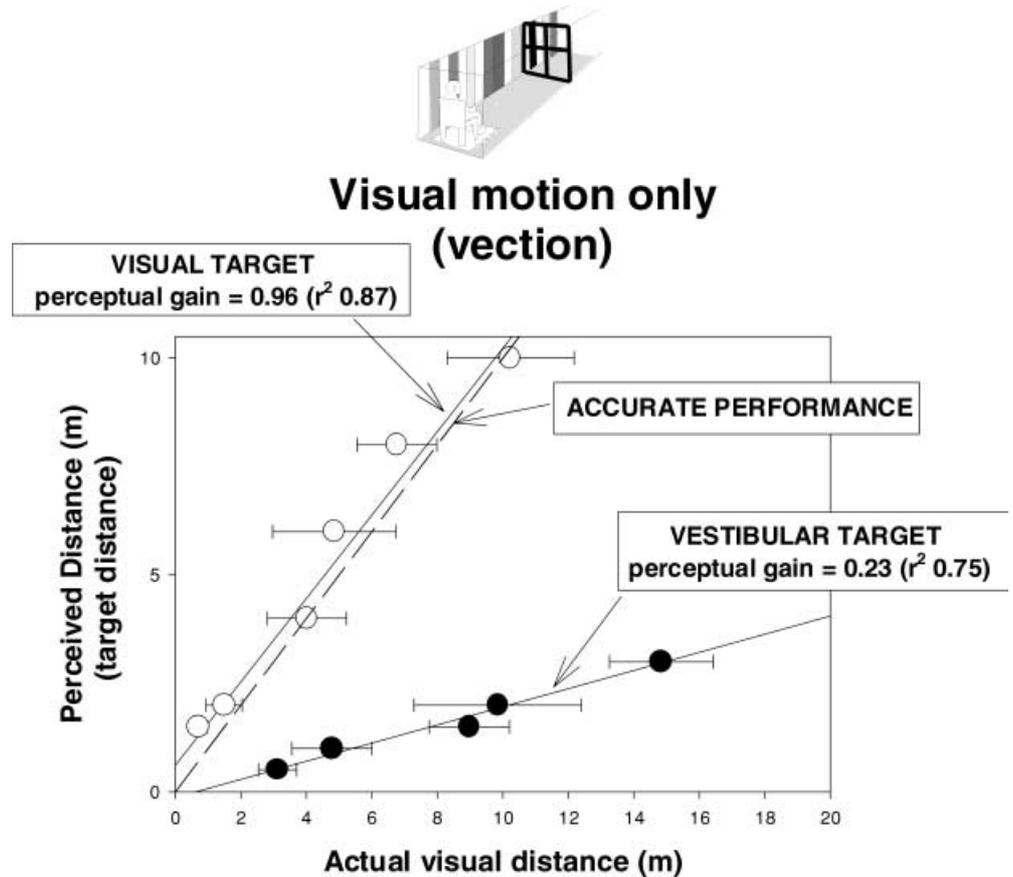


**Visual motion only (vection)**

VISUAL TARGET
perceptual gain = 0.96 ($r^2$ 0.87)

ACCURATE PERFORMANCE

VESTIBULAR TARGET
perceptual gain = 0.23 ($r^2$ 0.75)

Perceived Distance (m) (target distance)

Actual visual distance (m)

Vection perceptually equivalent to a physically presented target distance

When subjects had the target distance physically demonstrated to them (see Fig. 2) and were then subjected to purely visual motion, inducing a sense of vection, they needed to travel much further than the target distance before they perceived they had travelled through that distance. This is shown by the filled symbols of Fig. 4. The regression line through these data indicates a perceptual gain (perceived/actual motion) of only 0.23 ($r^2$ 0.75).

Combined visual and physical motion cues perceptually equivalent to a visually presented target distance

The experiments described above each presented motion cues through only a single modality, either physical motion alone (in the dark) or visual motion alone (no physical motion). What happens when both cues are presented at the same time as they are under normal conditions? For these trials subjects were first shown a target distance visually. They then moved down the virtual corridor at the same time as they moved physically and had to indicate when they had travelled through the previously presented target distance. The visual and physical movements were linked such that a given physical motion could be paired with different visual motions. We compared conditions in which the visual motion was either twice, equal to or half the physical motion.

If the visual and physical motions are different then, at the point the subjects press the button to indicate they have arrived at the target distance, they have travelled one distance in visual terms but a different distance physically. We therefore analysed each of these aspects of the subjects' motion separately. The results are presented in Fig. 5. Figure 5A shows the visual movements that were judged perceptually equivalent to the visually presented target distances while Fig. 5B plots the physical movements that were judged perceptually equal to those same target distances. If the perceptions of distance travelled were determined by the visual aspect of the stimulus, then the perceived distance plotted in visual terms (Fig. 5A) should depend only on the visual motion (visual capture) and all the data should cluster around the visual perceptual gain of 1.0 corresponding to the perceptual gain recorded when only visual cues were available (Fig. 4 open circles). The visual/visual perceptual gain line is shown superimposed on Fig. 5A. If, instead, the perceptual distance were determined only by the vestibular aspect (vestibular capture), then the data when plotted in terms of the physical motion should cluster around the perceptual gain of 2.07 corresponding to the gain recorded when only non-visual information was

**Visual and physical motion at the same time**

**Matched to visually presented target**

**A** Perceptual gain 0.96 matching visual motion to a visual target (from Fig 4)

**B** Perceptual gain 2.07 matching physical motion to a visual target (from Fig 3)

● Vision = 0.5 x physical
○ Vision = physical
▼ Vision = 2 x physical

**Visual and physical motion at the same time**

**Matched to physically presented target**

**A** accurate performance / Perceptual gain 0.23 matching visual motion to a physical target (from Fig 4)

**B** accurate performance / Perceptual gain 1.04 matching physical motion to a physical target (from Fig 3)

▲ vision = 2 x physical
○ vision = physical
● vision = 0.5 x physical

**Fig. 5A, B** Judgements of the distance of self motion matched to a visually presented target when physical displacement and visual displacement were different. Visual motion was either two times (*filled triangles*), equal to (*open circles*) or half (*filled circles*) of the simultaneous physical motion. Perceived distance (vertical axis, target distance) is plotted against the visual distance at which subjects indicated they had travelled through the target distance in **A** and against the corresponding physical distance in **B**. Data are the average of 12 subjects' responses with standard errors. The *dashed lines* indicate a perceptual gain of unity. The *solid line* in **A** shows the visual perceptual gain obtained by matching visual movement to a visually presented target distance (0.96 from Fig. 4). The *solid line* in **B** shows the vestibular perceptual gain obtained by matching physical movement to a visually presented target distance (2.7 from Fig. 3). The data cluster more closely to the vestibular perceptual gain when plotted against physical movement in **B** than they do to the visual perceptual gain when plotted against visual movement in **A**. This indicates that the physical motion component dominates the determination of the perceptual equivalence to a particular visually presented target irrespective of the concomitant visual motion

**Fig. 6A, B** Judgements of the distance of self motion matched to a physically presented target when physical displacement and visual displacement were different. Format as for Fig. 5. The *dashed lines* indicate a perceptual gain of unity. Visual motion could be either half (*filled circles*), equal to (*open circles*) or twice (*filled triangles*) the physical motion. As for the response to visually presented targets shown in Fig. 5, the data almost superimpose when plotted in vestibular terms. This indicates that the physical motion component dominates in the determination of the perceptual equivalence to a particular physically presented target, irrespective of the concomitant visual motion

**Combined visual and physical motion cues perceptually equivalent to a physically presented target distance**

Subjects were asked to match combinations of visual and physical motion to a physically demonstrated distance. The visual and physical movements were linked such that a given physical motion could be paired with different visual motions. As for experiments with visually presented distances, we compared conditions in which the visual motion was either twice, equal to or half the physical motion. Data are plotted in Fig. 6 using the same format as Fig. 5. The data are plotted in terms of the visual position in the virtual corridor in Fig. 6A and again in terms of the physical displacement at the time of the button press in Fig. 6B. The data clustered when plotted in terms of the physical motion required, but spread out when plotted in visual

available (Fig. 3 open and purple circles). This line is shown superimposed on Fig. 5B.

Figure 5 shows that, far from clustering around a perceptual gain of 1.0, the data are spread out when plotted in visual terms (Fig. 5A). When plotted in terms of the simultaneous physical motion, however (Fig. 5B), the data do cluster with an average perceptual gain (perceived/actual motion) of 4.54 ($r^2$ 0.78). The perceptual gains for each condition are given in Table 1.

terms. The perceptual gains (perceived/actual motion) are summarised in Table 1.

## Discussion

The accuracy of people's perception of how far they have moved needs to be defined functionally. Although people's established habit of reporting distances relative to a yardstick encourages us to believe that there are absolute measures of length such as miles and metres, these are in fact relative measures. A distorted perception of a metre, for example, cannot be expressed in metres. The experiments described in this paper used the subject's self motion itself as a direct measure of perceptual equivalence. By reporting when they have travelled through a certain perceptual distance (i.e. a previously presented target distance), subjects indicate that the actual motion they were exposed to was perceptually equivalent to this distance. Subjects could not perform any of our tasks using a remembered duration. For visually presented targets there was no temporal cue and for physically presented distances, unpredictable accelerations were used for the subsequent movement. Our results indicate that humans are able to make consistent estimates of the distance of their self displacement using either visual (Fig. 4) or physical motion cues (Fig. 3), at least when movement is at a constant acceleration.

Specifically our experiments have demonstrated that for constant acceleration movement of between 0.1 and 0.3 m.s$^{-2}$: (1) the perceived distance of motion evoked by optic flow was accurate relative to a visual target (Fig. 4 open circles) but (2) was perceptually equivalent to about half the physical motion (Fig. 3 open and purple circles), (3) the perceived distance of physical motion in the dark was accurate relative to a previously presented physical motion (Fig. 3 red, green, yellow and black symbols) but (4) was perceptually equivalent to about four times the visual motion (Fig. 4 filled circles), and (5) the perceived distance of self motion when both visual and physical cues were present in different amounts was more closely perceptually equivalent to the physical motion experienced and not the simultaneous visual motion even when the target was presented visually (Figs. 5, 6). These perceptual equivalences are summarised in Table 1.

### Comparison with other studies

Other studies comparing physical motion with physically presented target distances have confirmed that such motions can be accurately matched (Berthoz et al. 1995; Grasso et al. 1999) and that physical motion can lead to the generation of accurate eye movements (Israël and Berthoz 1989). Experiments involving varying the motion profile between sample and test (Israël et al. 1997) or requiring the construction of a spatial map to solve a task (Mittelstaedt 1980, 1999; Loomis et al. 1999) have showed that position information can be derived from a variety of self motion information.

Israël et al. (1993) matched a visually presented target distance with physical motion over short distances and found that subjects needed less physical motion (0.24 m) to match a visual distance (0.8 m). This overestimation, by a factor of between 3 and 5 for acceleration values around 0.5 m.s$^{-2}$, was also found when subjects were asked to estimate displacement in metres (Golding and Benson 1993) perhaps reflecting a visualised comparison. The present study reports overestimates of between two times for vestibular matches to visual targets (Fig. 3) and 4.3 times for visual matches to vestibular targets (reciprocal of perceptual gain: Fig. 4) which are compatible with these values. The overestimation of self motion using physical cues has also been reported for motion in the z-axis (Young and Markmiller 1996) and under active motion conditions (Loomis et al. 1993).

Pavard and Berthoz (1977) demonstrated that visual motion sensitivity could be reduced during combined visual and physical motion. The reduction in the use of visual movement cues under our vision-plus-vestibular condition might represent another example of this phenomenon. Anecdotally, a number of our subjects reported that the visual motion appeared too slow. It is probable that the vibration of our cart (Seidman and Paige 1998) and other factors such as noise, wind and expectancies also might have contributed to our subjects' sensation of motion. Our aim in this study was not to define the mechanisms of the processing of non-visual cues but to contrast them with the effectiveness of visual-only cues.

### Using a head-mounted display

The experiments we report here used a head-mounted display and so-called 'virtual reality' or 'immersive' technology. How generalisable are our results for describing the processing of optic flow in general? Head-mounted displays have been shown to evoke eye movement and perceptual responses very comparable to those evoked by 'natural' stimuli (for example Kramer et al. 1998) and our head-mounted display evoked a powerful and appropriate sensation of self motion (vection). Most vision research simulates some selected aspect of the visual world out of its natural context. Thus, in a sense, most vision research uses 'virtual' reality rather than 'normal' or 'natural' vision. For example, work concerning linear optic flow and self motion has traditionally used large fields of moving dots presented on a fronto-parallel plane as its 'virtual reality' display (for example Duffy 1998). The present work used optic flow stimuli presented on a screen that subtended a similar extent (84°×65°) to conventional presentation methods and that was viewed at optical infinity. The advantage of using a head-mounted display slaved to head movement is that a complete visual surround can be simulated, leading to a sensation of a real, three-dimensional visual environ-

ment. Although these environments are not yet photographic in their detail, they allow controlled presentation of simple stimuli in a way comparable to conventional CRT or projection system displays, but they offer the added advantage of being able to dissociate visual from non-visual cues and thus represent an important step in bringing the investigation of real world, multisensory experience into the realm of controlled, scientific investigation.

The experiments which showed that matching real world target distances to physical motion in the dark was indistinguishable to matching virtual visual targets to the same motion are strong support for the validity of using a head-mounted display to present visual stimuli. Especially since in both cases the same unexpected phenomenon was observed in which too-short physical distances corresponded to the visually presented target distances.

Comparison of perceptual observations
with neural processing

Cells in the parietal cortex and hippocampus have been strongly implicated in processing self motion. Cells in these regions respond to the visual optic flow pattern associated with linear movement (parietal reviewed in Andersen 1997; Andersen et al. 1997, hippocampus reviewed in Taube 1998). Despite its close association with self motion and neural representation, however, optic flow is a dangerous cue to rely on in isolation when determining self displacement. The preliminary task of parsing the visual array into areas that might contain useful optic flow (earth-fixed targets reasonably close) as distinct from areas that provide either no information (distant objects) or confusing information (items moving with the viewer) is itself not trivial. Even when useful flow is identified, without a **scale** it is impossible to distinguish motions as different as interstellar travel and ordinary locomotion using optic flow alone. Optic flow can only be meaningfully used in the context of other sensory or perceptual information.

Until recently most studies of the vestibular contribution to the properties of the parietal cortex and hippocampus have concentrated on *angular* movement (for example, parietal: Thier and Erickson 1992; Brochtie et al. 1995 but see Grüsser et al. 1990, hippocampus: Sharp et al. 1995). However, studies are now emerging that report responses in the medial superior temporal area to both visual and physical linear motion (Duffy 1998; Bremmer et al. 1999). These studies form a much-needed completion to the extensive research implicating the parietal cortex in self motion processing.

Linear-motion sensitive visual-otolithic convergence has been found in the vestibular nucleus (Daunton and Thomsen 1979) where there is a dominance of vestibular input for the higher frequencies of linear accelerations (Xerri et al. 1988) compatible with the vestibular dominance observed at high frequencies in postural control (Lestienne et al. 1977). The present psychophysical study is the first to provide a perceptual correlate of vestibular dominance during self motion and predicts that vestibular-related activity in the parietal cortex and/or hippocampus will be particularly pronounced during passive, linear motion.

Capture of the perception of linear self motion
by non-visual cues

When senses disagree one potential solution to the ambiguity is for one sense to be trusted more than another and hence to dominate the perception. Sometimes the domination is so complete that information from the subordinate sense, even though different from the dominant sense, *appears* to agree with the information from that sense. This illusory agreement is called *sensory capture*. A classic example of sensory capture is exploited by ventriloquism. In ventriloquism, vision is the capturing sense and sounds which actually arise from a different location are perceived as coming from the visually determined direction even though the auditory cues indicate it is coming from elsewhere. What we have described here is a very unusual example of intermodal sensory capture. First, it is unusual to show the visual system being dominated by any other sense in humans. Second, it is unusual because in previous examples of capture, a subject had two senses that gave different information. Here we have an example of one sense capturing another even when both could, theoretically, indicate the same thing: that the person has travelled a certain distance. The capturing phenomenon when the visual and physical motions were equal was only revealed by the fact that judgements of self motion in the dark are in such glaring error.

Although estimates of displacement can be obtained from physical motion cues, under the passive conditions of our experiment these cues are interpreted cautiously to overestimate self displacement. This conservative physical-cue generated estimate can dominate simultaneous optic flow cues underscoring the unreliability of that cue and providing a biological safety device by over-reacting to unexpected passive movement.

## References

Andersen RA (1997) Multimodal integration for the representation of space in the posterior parietal cortex. Philos Trans R Soc Lond B Biol Sci 352:1421–1428

Andersen RA, Snyder LH, Bradley DC, Xing J (1997) Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Annu Rev Neurosci 20:303–330

Berthoz A, Israël I, Georges-Francois P, Grasso R, Tsuzuku T (1995) Spatial memory of body linear displacement: what is being stored? Science 269:95–98

Bremmer F, Lappe M (1999) The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. Exp Brain Res 127:33–42

Bremmer F, Kubischik M, Pekel M, Lappe M, Hoffmann KP (1999) Linear vestibular self-motion signals in monkey medial superior temporal area. Ann N Y Acad Sci 871:272–81

Brochtie PR, Andersen RA, Snyder LH, Goodman SJ (1995) Head position signals used by parietal neurones to encode locations of visual stimuli. Nature 375:232–235

Daunton N, Thomsen D (1979) Visual modulation of otolith-dependent units in cat vestibular nuclei. Exp Brain Res 37:173–176

Duffy CJ (1998) MST neurons respond to optic flow and translational movement. J Neurophysiol 80:1816–1827

Esch HE, Burns JE (1995) Honeybees use optic flow to measure the distance of a food source. Naturwissenschaften 82:38–40

Esch HE, Burns JE (1996) Distance estimation by foraging honeybees. J Exp Biol 199:155–162

Gibson JJ (1950) The perception of the visual world. Houton Mifflin, Boston

Glasauer S, Amorim MA, Vitte E, Berthoz A (1994) Goal-directed linear locomotion in normal and labyrinthine-defective subjects. Exp Brain Res 98:323–335

Golding JF, Benson AJ (1993) Perceptual scaling of whole-body low frequency linear oscillatory motion. Aviat Space Environ Med 64:636–640

Grasso R, Glasauer S, Georges-Francois P, Israël I (1999) Replication of passive whole body linear displacements from inertial cues. Ann N Y Acad Sci 871:345–366

Grüsser O-J, Pause M, Schreiter U (1990) Localization and responses of neurons in the parieto-insular vestibular cortex of awake monkeys (Macaca-Fascicularis). J Physiol (Lond) 430:537–557

Harris LR, Jenkin M (1996) Comparing judgments of linear displacement using visual and vestibular cues. Invest Ophthalmol Vis Sci 37:2375–2375

Harris LR, Morgan MJ, Still AW (1981) Moving and the motion after-effect. Nature 293:139–141

Israël I, Berthoz A (1989) Contribution of the otoliths to the calculation of linear displacement. J Neurophysiol 62:247–263

Israël I, Chapuis N, Glasauer S, Charade O, Berthoz A (1993) Estimation of passive horizontal linear-whole-body displacement in humans. J Neurophysiol 70:1270–1273

Israël I, Grasso R, Georges-Francois P, Tsuzuku T, Berthoz A (1997) Spatial memory and path integration studied by self-driven passive linear displacement. I. Basic properties. J Neurophysiol 77:3180–3192

Koenderink JJ, Doorn AJ van (1975) Invariant properties of the motion parallax due to movement of rigid bodies relative to an observer. Opt Acta 22:773–791

Kramer PD, Roberts DC, Shelhamer M, Zee DS (1998) A versatile stereoscopic visual display system for vestibular and oculomotor research. J Vestib Res 8:363–379

Lappe M (2000) Neuronal processing of optic flow. Academic Press, San Diego

Lappe M, Rauschecker JP (1994) Heading detection from optic flow. Nature 369:712–713

Lestienne F, Soechting JF, Berthoz A (1977) Postural readjustments induced by linear motion of visual scenes. Exp Brain Res 28:363–384

Loomis JM, Klatzky RL, Golledge RG, Cicinelli JG, Pellegrino JW, Fry PA (1993) Nonvisual navigation by blind and sighted: assessment of path integration ability. J Exp Psychol Gen 122:73–91

Loomis JM, Klatzky RL, Golledge RG, Philbeck JW (1999) Human navigation by path integration. In: Golledge RG (ed) Wayfinding, mapping and spatial behavior. Hopkins Press, Baltimore, pp 125–152

Lowenstein OE (1974) Comparative morphology and physiology. In: Kornhuber HH (ed) Handbook of sensory physiology. The vestibular system. Springer, Berlin Heidelberg New York, pp 75–124

Mayne R (1974) A systems concept of the vestibular organs. In: Kornhuber HH (ed) Handbook of sensory physiology. The vestibular system. Springer, Berlin Heidelberg New York, pp 493–580

Mittelstaedt H (1980) Homing by path integration in a mammal. Naturwissenschaften 67:566–567

Mittelstaedt H (1997) Interaction of eye-, head-, and trunk-bound information in spatial perception and control. J Vestib Res 7:283–302

Mittelstaedt H (1999) The Role of the otoliths in perception of the vertical and in path integration. Ann N Y Acad Sci 871:334–344

Ohmi M (1996) Egocentric perception through interaction among many sensory systems. Cognitive Brain Res 5:87–96

Parker DE, Wood DL, Gulledge WL, Goodrich RL (1979) Self-motion magnitude estimation during linear oscillation: changes with head orientation and following fatigue. Aviat Space Environ Med 50:1112–1121

Pavard B, Berthoz A (1977) Linear acceleration modifies the perception of a moving visual scene. Perception 6:529–540

Redlick FP, Harris LR, Jenkin M (1999) Active motion reduces the perceived self displacement created by optic flow. Invest Ophthalmol Vis Sci 40:4199.2

Royden CS, Banks MS, Crowell JA (1992) The perception of heading during eye-movements. Nature 360:583–587

Seidman SH, Paige GD (1998) Perception of translational motion in the absence of non-otolith cues. Soc Neurosci Abstracts 24:162.10

Sharp PE, Blair HT, Etkin D, Tzanetos DB (1995) Influences of vestibular and visual-motion information on the spatial firing patterns of hippocampal place cells. J Neurosci 15:173–189

Smith PF (1997) Vestibular-hippocampal interactions. Hippocampus 7:465–471

Srinivasan MV, Zang S, Bidwell N (1997) Visually mediated odometry in honeybees. J Exp Biol 200:2513–2522

Srinivasan MV, Zhang S, Altwein M, Tautz J (2000) Honeybee navigation: nature and calibration of the "odometer". Science 5454:851–853

Stevens SS (1955) The measurement of loudness. J Acoust Soc Am 27:815–829

Taube JS (1998) Head direction cells and the neurophysiological basis for a sense of direction. Prog Neurobiol 55:225–256

Telford L, Howard IP, Ohmi M (1995) Heading judgements during active and passive self-motion. Exp Brain Res 104:502–510

Thier P, Erickson RG (1992) Vestibular input to visual-tracking neurons in area MST of awake rhesus monkeys. Ann N Y Acad Sci 656:960–963

Vanni-Mercier G, Magnin M (1982) Single neuron activity related to natural vestibular stimulation in the cat's visual cortex. Exp Brain Res 45:451–455

Warren WH, Morris MW, Kalish M (1988) Perception of translation heading from optical flow. J Exp Psychol Hum Percept Perform 14:646–660

Warren WH, Blackwell AW, Kurtz KJ, Hatsopoulos NG, Kalish ML (1991) On the sufficiency of the velocity field for perception of heading. Biol Cybern 65:311–320

Xerri C, Barthelemy J, Borel L, Lacour M (1988) Neuronal coding of linear motion in the vestibular nuclei of the alert cat. 3. Dynamic characteristics of visual-otolith interactions. Exp Brain Res 70:299–309

Young LR, Markmiller M (1996) Estimating linear translation: saccular versus utricular influences. J Vestib Res 6:S13

Zikovitz DC, Harris LR, Jenkin M, Kreichman D (1998) Vestibular dominance in estimating the distance of self motion: misaligned optic flow and vestibular cues. Invest Ophthalmol Vis Sci 39:5063