

# Vestibular cues and virtual environments: choosing the magnitude of the vestibular cue

Laurence Harris<sup>1,3</sup> Michael Jenkin<sup>2,3</sup> Daniel C. Zikovitz<sup>3</sup>  
Departments of Psychology<sup>1</sup>, Computer Science<sup>2</sup>, and Biology<sup>3</sup>  
York University  
4700 Keele St., Toronto, Ontario, Canada, M3J 1P3

## Abstract

*The design of virtual environments usually concentrates on constructing a realistic visual simulation and ignores the non-visual cues normally associated with moving through an environment. The lack of the normal complement of cues may contribute to cybersickness and may affect operator performance. In [3] we described the effect of adding vestibular cues during passive linear motion and showed an unexpected dominance of the vestibular cue in determining the magnitude of the perceived motion. Here we vary the relative magnitude of the visual and vestibular cues and describe a simple linear summation model that predicts the resulting perceived magnitude of motion. The model suggests that designers of virtual reality displays should add vestibular information in a ratio of one to four with the visual motion to obtain convincing and accurate performance.*

## 1 Introduction

Virtual environments have been proposed for many varied applications including the treatment of phobias[12], providing control for a mobile robot[1] and even providing a safe place for children to practice crossing the street[4]. In many virtual simulations effort has concentrated on constructing an accurate, realistic, visual simulation while other environmental and motion cues tend to be ignored.

A fundamental aim in the design of virtual reality (VR) systems is to give people the convincing impression that they are moving around within a simulated world. But the emphasis on the visual display means that many of the normal cues to motion are not being provided. A virtual environment with only visual cues ignores for example the physical cues to linear self motion which are normally signalled by the acceleration-sensitive otolith division of the vestibular

system[7]. Periods of constant velocity cannot be registered by this acceleration-sensitive system; but for motions with changing velocities, position can be obtained by double integration of the acceleration signal. Humans are able to use vestibular information to assess a position change [8, 5, 2, 6] and direction of travel [14, 13, 9].

So when an operator “moves” within a vision-only virtual environment they are presented with at least two contradictory cues to their motion; the visual system carefully arranged to signal motion of the operator while the vestibular system provides a signal which is consistent with no motion or motion at constant velocity. Disparate cues generate an inter-sensory conflict and this has been postulated as a major reason for motion sickness [11, 10]. The associated nausea undoubtedly reduces the efficiency of an operator working in a virtual environment.

All of these points argue strongly for the addition of real-motion cues. But a question that has daunted designers of VR systems that try and include vestibular cues is what vestibular cues to apply? This paper describes an experiment in which we presented optic flow and non-visual cues in various proportions. We found that the addition of vestibular cues has a very powerful influence on subject’s perceptual experience: their judgements of self motion were critically affected by the amount of physical motion that accompanied their visual motion. The effect could be convincingly modelled by a simple weighted linear summation of visual and vestibular information.

## 2 Methods

### 2.1 General

Different magnitudes of visual and vestibular cues were presented by moving subjects physically at a constant acceleration as they sat on a cart wearing a vir-

tual reality helmet which put them in a virtual corridor. This allows various combinations of visual and non-visual cues to be presented. The general arrangement of the apparatus we used is shown in Figure 1. Optic flow was presented during physical motion at constant acceleration. The optic flow was compatible with movement of accelerations that differed systematically from the physical motion by a factor of a quarter to four times. Subjects judged the distance they felt they had travelled along the virtual corridor.

## 2.2 Subjects

Seventeen subjects took part in these experiments. They had normal visual acuity and no history of vestibular or balance problems. Experiments were approved by the York Ethics Approval Committee. Subjects were paid for their participation at standard York subject rates.

## 2.3 The cart

Subjects sat on a chair mounted on a cart (Figure 1). The cart could physically move through a distance of up to four metres. It was mounted on low-friction, in-line skate wheels and ran on a smooth floor. The cart was attached by a rope to a weight. The weight could be dropped through a distance of 1.5 metres which, by an arrangement of pulleys, pulled the cart through a distance of up to 4.5 metres. When the weight was released it fell with a downward force of mass  $\times$  g and pulled the cart at a constant acceleration ( $0.1 - 0.4 \text{ m.s}^{-2}$  depending on the relationship between the mass of the weight and the mass of the subject). Cart position was transduced by running a thin, earth-fixed wire around the optical-encoder shaft of a mouse that was mounted on the cart. This 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. The calibrated signal from the mouse was then sampled by the computer and stored.

## 2.4 Visual display

Subjects viewed a  $84^\circ \times 65^\circ$  display presented on a single-screen, Liquid Image MRG3 virtual reality helmet equipped with a six degrees-of-freedom Flock of Birds head tracker. The reference transmitter for the Flock of Birds was mounted on the cart and therefore the head tracker reported the position of the head relative to the cart. The display simulated a virtual corridor 50m long, 2m wide and 2.5 m high, based

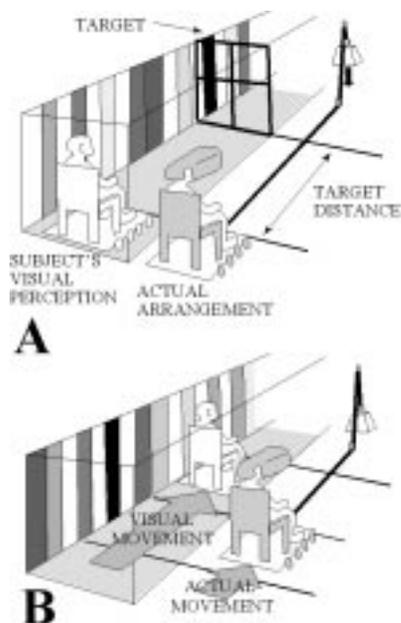


Figure 1: *Physical motion equipment.* This figure shows experimental apparatus we used for varying the relationship between visual and vestibular cues. The left side of each panel shows the subject's visual perception and the right side shows the actual motion for the visual plus vestibular condition. Subjects 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 along a 4 meter track. Subjects wore a virtual reality helmet (see text) which put them in a virtual corridor. Panel A shows the starting conditions. Note the target presented in the corridor. Panel B shows typical movements used in these experiments. The target is not present during subjects' movement. Subjects indicated when they had reached the target's original position. In the example shown, the physical motion is half that of the visual motion. For vision-only trials, the chair did not move but the subject's visual perception was also as in panel B. Dark conditions (vestibular only) were also run in which the display was blanked during the movement.

on the dimensions of a typical corridor at York University. The walls of the corridor were painted with vertical stripes 0.5m wide which changed colour on a random schedule approximately once a second. The flickering colour reduced the possibility that subjects merely tracked a stripe on the wall that lined up with the target. The floor and ceiling were blank. The visual position of the subject in the virtual corridor was derived from the vector sum of the physical position of the cart relative to the room, measured by the mouse (see above) and their head position relative to the cart, measured by the flock of birds. Vision-only displays were created by recording the physical motion of the cart and playing it back later to drive the virtual reality display. This ensured that the visual conditions were identical in the vision-only and the vision-plus-vestibular conditions.

## 2.5 Calibration of the visual display

It is important that perceived scale and distances in the visual display were correctly calibrated to the real world. To calibrate the optics of the VR display, subjects were presented with a target at a simulated distance (eg. two metres) and then lifted the VR helmet and viewed a real-world target at the same distance. Subjects then raised and lowered the helmet while the 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 then checked at several distances.

## 2.6 Visual motion of a different magnitude from the physical motion

The advantage of using a virtual reality visual display is that the visual displacement does not have to be of the same amplitude as the actual displacement of the cart. We describe the relationship between the subject's visual and physical position by the term "visual gain". Visual gain is defined as the ratio between the visual movement and the physical movement that drove it. The motion illustrated in Figure 1, for example has a visual gain of two indicating that the cart motion was multiplied by two before driving the visual display. Thus physical motion of one metre moved the subject two metres down the virtual corridor. A visual gain of unity corresponds to the 'natural condition'.

## 2.7 Procedure

In order to obtain a measure of how far subjects thought they had travelled under our various conditions, they were first given a reference distance and subsequently asked to indicate when they had moved through that distance. The reference distance could be presented either visually or vestibularly (see below). When this had been done, subjects indicated when they were ready and constant-acceleration movement commenced. One of three conditions was presented in a pseudorandom order. The three conditions were: (i) physical motion in the dark (physical motion only) (ii) visual motion with no physical motion (visual motion only) or (iii) both physical and visual motion; either the same or different magnitudes. Subjects were instructed to push a button to indicate when they felt themselves to have travelled through the reference distance.

## 2.8 Presentation of visual reference distance

To present the target distance visually subjects were presented with a target at a known position of between 0.5 and 10m in the virtual corridor. The target was a red frame that went all round the edges of the corridor with vertical and horizontal cross-pieces forming a cross (see Figure 1). The position of the target thus indicated a distance which subjects were asked to remember. 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. Targets presented in virtual reality were carefully calibrated against real targets at the same distance in the outside world. Targets were extinguished before the movement phase of the experiment commenced so there was no target present during either visual or physical movement.

## 3 Results

### 3.1 Judging the amplitude of physical motion in the dark

The present results confirm our previous studies ([3]) that established the response to motion in the dark under the conditions of these experiments. Subjects consistently and dramatically over-estimated their movement and pressed the button to indicate they had arrived at the target after having travelled through only about a third of the target distance. This

visual gain	Displacement from visual-only condition				p	Displacement from vestibular-only condition			
	judged equal to one meter	F*	dof			judged equal to one meter	F*	dof	p
0.25	0.13	5.2	62	p<0.05	0.52	0.39	107	ns	
0.5	0.19	46.2	73	p<0.01	0.37	0.18	118	ns	
1	0.45	148.2	106	p<0.01	0.45	0.27	151	ns	
2	0.62	140.4	121	p<0.01	0.31	3.7	166	ns	
4	0.95	3.2	81	ns	0.24	2.8	126	ns	
vis only	0.95								
vest only					0.30				

Table 1: *The slopes of Figures 2 and 3 indicate how many metres the optic flow and vestibular cues need to move in order to give the perception of moving a meter.*

is indicated in Figure 2, which plots distance actually travelled on the vertical axis, against the target distance by the line labelled “dark”. Physical motion was always forwards along the naso-occipital axis.

### 3.2 Judging the amplitude of motion from visual motion alone (vection)

Subjects were quite accurate at judging their motion under the vision only conditions. This is shown by the line labelled “vision only” in Figure 2 which has a slope of 0.95.

### 3.3 Response to combinations of visual and physical motion

Our experiments pulled apart the optic flow and vestibular cues by pairing a given movement down the virtual visual corridor with different actual motions of the cart along its track. The ratio between the visual movement and the vestibular movement is the visual gain. The visual gain is defined as:

$$\text{visual gain} = \text{visual motion} * \text{physical motion} \quad (1)$$

A visual gain greater than one indicates more visual movement than vestibular movement and a visual gain less than one indicates less visual movement than vestibular movement. Visual gain is expressed in diagrammatic form in Figure 2a. When the visual gain is not unity and subjects press the button to indicate they have arrived at the target distance, they have travelled a different distance in visual terms down the corridor than they have in vestibular terms across the floor. For example if subjects travel 1m across the floor with a visual gain of 4, they will have simultaneously travelled 4m down the visual corridor

(Figure 2a). Therefore we need to analyse the results separately in each of the visual and vestibular frames. The results are presented in Figure 2. Figure 2b shows the *visual* movements that were judged equivalent to the target distances and Figure 2c plots the vestibular movements that were judged equal to those same target distances. If the perception of distance travelled was decided by visual cues (visual capture), then the data obtained with any visual gain should fall on the ‘vision alone’ line of Figure 2b. If instead the perceptual distance was determined by vestibular cues (vestibular capture), then the data from different visual gains should lie on top of each other on the vestibular-only line of Figure 2c. The data are summarized in Table 1 and strongly support the vestibular capture hypothesis. All the curves except the visual gain of four are significantly different from the vision-only curve in Figure 2b and none of the curves is significantly different from the vestibular-only curves in Figure 2c. Clearly there is an overwhelming capture of the perception of the distance travelled by vestibular cues. But there is also an effect of vision: the lines do not superimpose exactly in vestibular terms. The slope when the visual gain was 0.25 is more than twice the slope when the visual gain was 4.0.

## 4 Discussion

### 4.1 Summary

When subjects were physically stationary and had only visual optic flow information on which to base their judgements their performance was accurate (“vision only” line in Figure 2b; slope = 0.95 see Table 1). When motion was in the dark (only vestibular

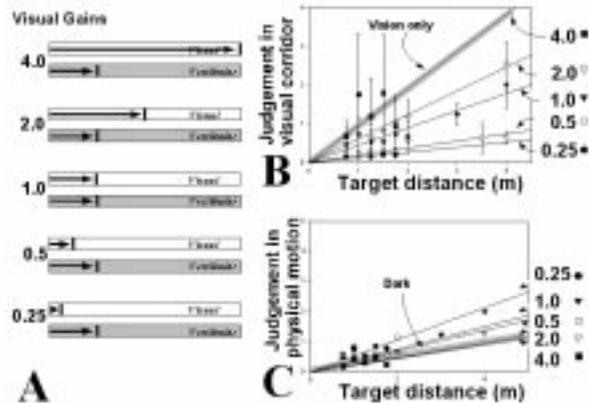


Figure 2: When visual and vestibular cues were of different magnitudes the cart’s physical displacement on the floor and the person’s visual displacement along the virtual corridor were different. This is shown diagrammatically in panel A in which the relative visual (open bars) and vestibular motions (shaded bars) are described as the ratio between the visual displacement and the actual displacement: the visual gain (see text). We used five visual gains from 0.25 to 4.0. The point at which subjects indicated they had travelled through the target distance is plotted in visual terms in panel B and the same data are replotted in vestibular terms in panel C. The horizontal axis shows the target distance. Data are the average of 17 subject’s responses with standard deviations shown in panel B. The thick line in panel B labelled “vision only” shows the performance when only visual motion was provided and the thick line in panel C labelled “dark” shows the performance during motion in the dark.

cues available; “dark” line in Figure 2) subjects consistently and dramatically over-estimated their movement and pressed the button much too early after having travelled through only about a third of the target distance. This phenomenon was found in response to either virtually-presented or real-world targets confirming our calibration procedure. When both visual and vestibular cues were presented, the vestibular cue dominated although an interaction effect could also be seen. No matter what the visual information, subjects considered they had reached the target when they had physically travelled through a particular distance. Their judgements were only slightly influenced by the amount of concurrent visual motion.

#### 4.2 Visual or vestibular terms?

When subjects are presented with different forwards accelerations through their visual and vestibular systems then, at the point they press the button to indicate they have arrived at a target’s location, their position in the visual corridor is different from their position along the floor. That is to say they have moved a different distance visually than they have physically. When the data were plotted in vi-

sual (Figure 2b) and in vestibular (Figure 2c) terms, they confirmed the result in our earlier study [3] that the vestibular system was much more influential in determining the point to push the button. But there is clearly an influence of vision. That is the vision plus vestibular curves do not **exactly** superimpose on the dark (vestibular only) data.

In order to quantify the visual component, the data were fitted to the following equation:

$$response = \frac{(vis * vis\ weighting) + (vest * vest\ weighting) + (interaction * int\ weighting)}{vis * vis\ weighting + vest * vest\ weighting + interaction * int\ weighting} \quad (2)$$

where: *response* is the position in the visual corridor at which subjects pressed the button, *vis* is the *visual component* which is equal to the *vis only slope\*target distance*, *vest* is the *vestibular component* which is equal to the *vest only slope\*target distance\*visual gain*, and *interaction* is the *visual component \* vestibular component*.

This model produced the following weightings: *vis* = 0.14, *vest* = 0.83, and *interaction* = 0.04.

Since the interaction term is very small compared to the others, we postulate the simple model shown in Figure 3 in which the visual and vestibular estimates

visual gain	visual displacement		vestibular displacement	
	judged equal to one meter	prediction from model	judged equal to one meter	prediction from model
0.25	0.13	0.21	0.52	0.85
0.5	0.19	0.28	0.37	0.55
1	0.45	0.40	0.45	0.40
2	0.62	0.65	0.31	0.32
4	0.95	1.15	0.24	0.29
vis only	0.95			
vest only			0.30	

Table 2: The prediction from our simple linear summation model (Figure 3) of visual and vestibular displacements that should be judged equal to one meter with a weighting of vision= 0.14 and vestibular = 0.83.

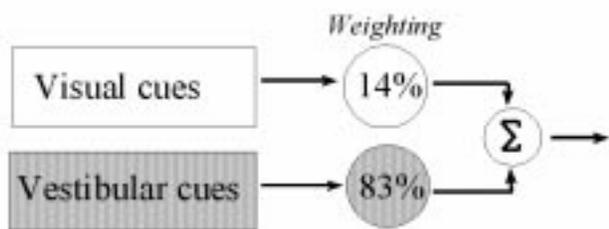


Figure 3: A simple model of weighted linear summation of visual and vestibular estimates of self motion.

of the motion are weighted and summed. The output of this model is summarized in Table 2 and in Figure 4 where it is compared to the data.

When plotted in vestibular terms (Figure 4 right hand side) the slopes do not vary very much with visual information because of the relatively small weighting given to vision. In order to bring the performance in visual terms back to the level of vision-only performance the visual gain needs to be close to four. In other words our data show that in order both to provide vestibular information and obtain accuracy in visual terms during passive motion, a virtual reality systems designer needs to provide about four times as much visual motion as vestibular motion.

At first glance our findings are rather counter-intuitive: why when visual and vestibular cues are both present are subjects not accurate at judging the distance they travel? Why is so much weight given to the the interpretation of the vestibular stimulus? It is important to remember that our experiments involve only two of the many cues normally available to help in the estimation of self motion. Most of the cues found during natural, active movement are missing, including muscle proprioception and a copy of motor

signals. Our visual cues were deliberately not as rich as the natural visual cues available during self motion. And our vestibular cue was particularly strong: continuous, passive acceleration is not normally encountered.

It seems probable to us that the over-active vestibular system we see working under our experimental conditions would be much toned down by active motion when normal motor commands and expectancies are present. Our experiments indicate that the vestibular system, like other proprioceptive systems, is most strongly stimulated by differences between expected and actual motion. This is how muscle spindles work, for example. Muscle spindles do not signal absolute muscle length but only differences between the actual and expected lengths. This is achieved by regulation at the level of the end organ by the gamma innervation to the contractile elements of the muscle spindle itself. In the case of the vestibular system, the comparison is probably carried out more centrally.

We can conclude from these experiments that, under our conditions at least, optic flow is not the dominant factor in determining the perception of a travelled distance. Although visual information can be used accurately when there are no competing cues, it is dominated by any concurrent vestibular stimulation. Over-interpretation of the vestibular stimulation causes people to feel they have travelled much further than they really have. Visual information has to be presented at four times the speed to keep up coincidentally with this enormously exaggerated interpretation of passive vestibular stimulation. This is good news for designers hoping to add convincing vestibular information: they only have to add vestibular information one quarter the amount of the visual motion.

## PREDICTIONS OF MODEL PLOTTED THROUGH THE DATA

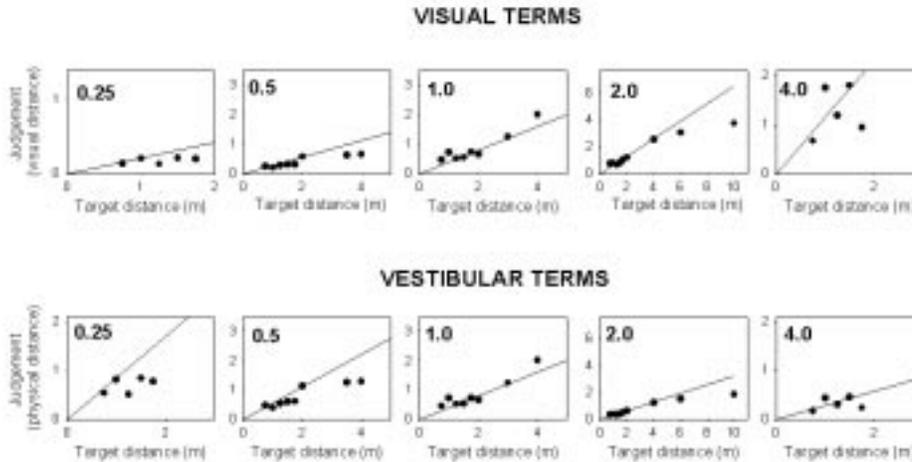


Figure 4: The output of the model of Figure 3 is plotted as a straight line in visual terms (top) and vestibular terms (bottom). The mean data for each visual gain are replotted from Figure 2 to compare with the model's performance.

### Acknowledgments

LRH and MJ are supported by a Collaborative Project Grant from the National Science and Engineering Research Council (NSERC) of Canada and the Centre for Research in Earth and Space Technology (CRESTech) of Ontario.

### References

- [1] N. Aucoin, O. Sandbekkhaug, and M. Jenkin. An immersive 3D user interface for mobile robot control. In *IASTED Int. Conf. on Applications of Control and Robotics*, pages 1–4, Orlando, FL, 1996.
- [2] A. Berthoz, I. Israel, P. Georges-Francois, R. Grasso, and T. Tsuzuku. Spatial memory of body linear displacement: what is being stored? *Science*, 269:95–98, 1995.
- [3] L. R. Harris, M. Jenkin, and D. C. Zikovitz. Vestibular cues and virtual environments. In *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS)*, pages 133–138, Atlanta, GA, 1998.
- [4] D. P. Inman, K. Loge, and J. Leavens. VR education and rehabilitation. *JACM*, pages 53–58, 1997.
- [5] I. Israel, N. Chapuis, S. Glasauer, O. Charade, and A. Berthoz. Estimation of passive horizontal linear-whole-body displacement in humans. *J. Neurophysiol.*, 70:1270–1273, 1993.
- [6] J. M. Loomis, R. L. Klatzky, R. G. Golledge, J. G. Cicinelli Pellegrino, J. W. Pellegrino, and P. A. Fry. Non-visual navigation by blind and sighted: assessment of path integration ability. *J. Expt. Psych. (General)*, 122:73–91, 1993.
- [7] O. E. Lowenstein. Comparative morphology and physiology. In H. H. Kornhuber, editor, *Handbook of Sensory Physiology, Vestibular System*, pages 75–124. Springer Verlag, New York, NY, 1974.
- [8] A. Mayne. A systems concept of the vestibular organs. In H. H. Kornhuber, editor, *Handbook of Sensory Physiology, Vestibular System*, pages 493–580. Springer Verlag, New York, NY, 1974.
- [9] M. Ohmi. Egocentric perception through interaction among many sensory systems. *Cog. Brain Res.*, 5:87–96, 1996.

- [10] C. Oman. Motion sickness - a synthesis and evaluation of the sensory conflict theory. *Canadian J. of Physio. and Pharmacol.*, 68:294–303, 1990.
- [11] J. T. Reason. Motion sickness adaptation: a neural mismatch model. *J. Roy. Soc. Med.*, 71:819–829, 1979.
- [12] D. Strickland, L. Hodges, M. North, and S. Weghorst. Overcoming phobias by virtual exposure. *JACM*, pages 34–39, 1997.
- [13] L. Telford, I. P. Howard, and M. Ohmi. Heading judgements during active and passive self-motion. *Expt. Brain Res.*, 104:502–510, 1995.
- [14] W. H. Warren, M.W. Morris, and M. Kalish. Perception of translation heading from optical flow. *J. Exp. Psych.: Human Percept. and Perf.*, 14:640–660, 1988.