Empirical investigations into the fragmentary nature of visual perception

Simon K. Rushton

Doctor of Philosophy
University of Edinburgh
1998

Declaration

I declare that this thesis was written by me and that the research reported herein was conducted by me, unless otherwise indicated.

September 1998.

Abstract

The proposition at the heart of this thesis is that the visual system may best be analysed as a collection of independent neural systems, each of which is dedicated to determining a specific environmental property or relation, through being attuned to only a subset of the information available in the optic array.

The first part of this proposition derives from the theory of Modularity which suggests that perception is realised through a number of functionally, and often anatomically, distinct neural systems specialised for specific functions (Fodor,1983). The second part derives from an interpretation of direct perception which suggests that the visual system determines a property or relation by use of only one of the multiple sources of information available (Cutting, 1986).

After a brief review of the tenets of Modularity I present the results of some empirical investigations which bring to the fore the limitations of Fodor's (1983) theory of modularity. The areas covered include: Synaesthesia, where modularity appears to break down or at least go soft at the edges; Brain injury in general and specifically unilateral visual neglect, the fracturing of the phenomenal world; Novel display devices that attempt to pull apart the two coupled-systems of accommodation and vergence.

The primary concern of the thesis is the "input" to modular systems. Specifically, whether modular systems are sensitive to just a restricted subset of the information potentially available. The main body concentrates on two key functions of visual perception, the timing of interception and the perceptual control of direction of locomotion.

I find first that interceptive timing does not rely on a sole source of information, optical looming. Rather it uses both looming and binocular disparity. A model, the *dipole* model of perception of time-to-contact (TTC) is described and tested. Second, I find that disparity is not readily ignored. Even when it is unreliable as a cue to time-to-contact it still retains its influence. This is compatible with models that involve the early combination of disparity and looming such as the *dipole* model.

Next, I switch to the perceptual control of locomotor direction. Following an empirical investigation, I conclude that perceptual control of locomotion does not exploit depth information. This conclusion stands in contradiction to a report based upon perceptual judgement of locomotor direction (van den Berg & Brenner, 1994). I then go on to report a study that undermines the widely held assumption that perceptual control of locomotion on foot relies on optic flow. It is found that walking is guided by a single source of information, perceived ego-centric direction. I conclude with a description of a complete theory of locomotion based upon use of ego-centric direction.

Implications of the main sequence of empirical studies for practical issues, including the design of computer generated virtual environments, and interventions for patients with unilateral visual neglect, are discussed during the course of the thesis

Acknowledgements

Julie Harris for providing the encouragement I needed and generous help when I asked. Hamish MacLeod, my supervisor, for trusting me with a very long length of rope. John Wann for putting me right as he is well used to doing. David Johnson, Tricia Riddell, Madeleine Grealy, Martin Smyth, John Bullinaria, Maugan Lloyd, David Lee, Mark Mon-Williams, Martin Corley, Trine Langaas, Anna Plooy and David Swapp who have all collaborated with me at some time or another. Julie Harris, Martin Smyth, John Wann and Kenneth Hamilton for proof reading and invaluable comments. The Psychology Technicians, both the Jimmy's and Davy for helping to keep me grounded and my experiments running.

Table of Contents

Note on contents and overview

Papers as chapters

With the exception of one chapter (chapter 3), all the work reported has already been written up as papers. Some of the papers are accepted (chapter 5), some are almost accepted (chapter 2 &4), and some are waiting to be submitted (chapter 6). I have not deliberately gone through and rewritten the work to obscure the origins of the chapters but rather have included a draft of the paper in each chapter. I have standardised the referencing and have sandwiched most of the papers with introductory material and sometimes supplemental material. In many cases I had added a foreword and/or afterword. In the forewords I provide some additional background and in the afterwords I try to draw out some of the implications of the results reported. The chapters are organised as a progression with a single switch from one sequence of papers to another around the middle of the manuscript.

Collaborators

Nearly all of the work I report in this thesis has been conducted with collaborators. The major reason for this is because my PhD has been done part-time PhD and I have not had the luxury of having research time entirely to myself to do with as I wish.

The main body of this thesis describes a subset of the research I have been involved in. I have selected the research reported carefully, I took the inclusion criterion to be papers on which I was first author. In some of the cases this means that the work is 95% my own, and in one case, at the other end of the scale, the *dipole* model paper reflects a 50/50 collaboration with John Wann (the model is mine, the results are joint, the formal derivations are John's and we both owe Martin Smyth and Anna Plooy thanks for having collected a second dataset for us) . I should highlight that I choose collaborators who are good at maths because maths is not my strong point: I solve problems in my head, graphically on bits of paper or through computational simulation. Therefore, I identify Julie Harris and John Wann as having turned my scribblings, sketches and arm-waving descriptions into formal derivations and algebra and thank them profusely for having done so.

The main chapters are concerned with the information used in the perception of time-to-contact (TTC) and direction of locomotion.

Chapter 1: Introduction and Background

In the first chapter, I briefly describe the theory of modularity and indicate some shortcomings of a simple statement of the theory. At the end of the chapter I present extended abstracts of research I have conducted that bear on the theory of modularity. Brief introductions to the perception of time-to-contact and the perception of direction of locomotion are included as background to the later empirical chapters.

Chapter 2: The *Dipole* model of perception of time-to-contact

Foreword: A brief note on the optical specification of time-to-contact and the background to the paper.

Paper: *Weighted Combination of Size and Disparity: A computational model of the timing of ball-catch.* This paper starts by revisiting a dataset described several years ago in Wann & Rushton (1995). When we initially published the data we were able to do little more than describe the results. This paper presents the *dipole* model that was inspired by the work and which was designed to account for the results of Wann & Rushton. The model is evaluated in a second experiment designed to provide an extreme test of the model. I came up with the model through running a series of empirical simulations, John Wann took my crude formalisation of the model and worked backwards to provide the derivations. The paper was jointly written by the two of us.

Afterword: Some thoughts about strong versus weak fusion and a flow implementation of the dipole model.

Chapter 3: Unreliable disparity and interceptive timing

Foreword: A brief note about "strong fusion" and the potential for adaptation in the TTC system.

Paper: *Can the influence of disparity be altered for interception in front of the face?* This dataset examines whether it is possible to reduce the influence of disparity on timing of interceptive actions. Observers attempt to intercept an approaching ball whilst wearing telestereoscopes or Cyclopean glasses. The data was collected several years ago with Mark Mon-Williams and John Wann. It was never analysed or published. I have written a lengthy appendix is included that details different routes to the estimation of TTC.

Afterword: -

Chapter 4: Depth and the perception of direction of locomotion

Foreword: Note about the paper to follow.

Paper: *Steering, Optic Flow and the respective importance of depth and retinal motion distribution.* A paper by van den Berg & Brenner suggested that depth information is used in the perception of heading. In this paper we revisit this matter with some novel manipulations and an active steering task. This experimental work was designed and conducted by myself primarily as material for this thesis. I was given input on the stereo manipulation by Julie Harris. However the results proved very interesting and so I wrote them up as a paper with Julie Harris and that paper is included here.

Afterword: A discussion, drawn from a paper by myself and Julie Harris on the importance of this finding for the design of VR systems.

Chapter 5: The influence of perceived location when walking

Foreword: A brief note on the history of the study.

Paper: *The role of perceived location during locomotion on foot.* It is commonly held that perception and control of locomotor direction is based upon use of the optical flow field. This paper re-examines this matter with a study involving observers wearing prism glasses. Maugan Lloyd collected the data for this paper for me for his final year project. The results were written into the paper included here with Julie Harris.

Afterword: A brief discussion of the results and how they may relate to the veering walks of patients with unilateral visual neglect and hemianopia. The discussion is based upon a paper I wrote on this matter.

Chapter 6: An eccentric model of walking

Foreword: Brief note about the paper

Paper: *An eccentric theory of control of locomotor direction: extending Llewellyn's (1971) model.* This paper was an attempt by myself to articulate a fuller and coherent model of control of locomotor direction when walking. Julie Harris collaborated on the paper providing the derivations of several equations and constructive feedback on the paper content.

Afterword: -

Chapter 7: Discussion and Conclusions

Conclusion: A restatement of the findings reported in the empirical chapters and a consideration of information and modularity.

Chapter 1:
Introduction and Background

1.1 Overview

The research presented in this thesis has the common theme of modularity. Although the issue of modularity is understated in the main body of the thesis, the research described can be best viewed against this background. I start with a brief description of modularity, to orient the reader, and then identify some areas of particular interest. I draw attention to the relevance of some of my research (which is not part of the main sequence of studies, but is reported *en précis* at the end of this chapter), to an appreciation of some of the issues associated with the theory of modularity. The primary concern of the research reported in this thesis is the "input" to modular systems. Two fundamental human behaviours are intercepting an approaching object and controlling direction of locomotion. It has been argued that the demands of the former (processing time and precision) dictate that perception of motion-in-depth and time-to-contact (TTC) must be processed by a specialised system (McLeod et al, 1985). A similar argument applies to the latter. Thus interceptive timing and perceptual control of locomotor direction make ideal subjects of study.

1.2 An introduction to Modularity

A module is defined as a cognitive system that is 'domain specific, innately specified, hardwired, autonomous and not assembled' (Fodor, 1983, p37). Modularity, the theory that perceptual systems (including language) are comprised of a collection of such modules has found common acceptance. In any modern textbook we can read that "V5" (or "MT") is an area of the visual cortex which is specialised for motion perception. The visual cortex being itself an area of the brain specialised for vision. If we take Zeki

(1993), "A Vision of the Brain", as an example, we are told that the history of neuroanatomy starts in 1861 in Paris with Broca discovering an area of the brain responsible for language. Modularity is certainly pervasive in modern psychology and neuroscience (see Shallice, 1988 for a discussion about neuropsychology and the theory of modularity).

The seminal modern text on modularity is Fodor (1983). Fodor made the case for modularity and enumerated certain key features of modular systems: dedicated neural architecture; mandatory operation; informational encapsulation; inaccessibility to central processes; rapid speed; shallow outputs; domain specificity; characteristic pattern of breakdown; fixed pattern of development.

Briefly considering these features: Dedicated neural architecture, rapid speed and domain specificity are all self-explanatory. Mandatory operation indicates that it is not possible to stop the functioning of the module. Only a subset of the information in the brain is available to a module, this is informational encapsulation. Inaccessibility to central process means one remains unaware of how a module processes an input. Shallow outputs means that a module delivers only basic categorisations. Characteristic pattern of breakdown - when a module is damaged there will be a specific impairment rather than just a general decrement in brain function. Fixed pattern of development denotes that development of a module is innately specified and proceeds in a characteristic sequence.

Fodor's theory appears to have a few shortcomings. These are highlighted by some of the research I describe below and "patched" with ideas taken from elsewhere. The issues discussed are: the independence of modules (from other modules, and central processes), plasticity and modularity, and the input to modules.

1.2.1 The degree of autonomy of modules

If modules are absolutely autonomous, the state of one module should have no influence on another unless the two are hierarchically linked. Is such a strict definition useful? Shallice (1988) argues that when considering neural systems we should not attempt to label them as modular or non-modular, but rather assess their degree of modularity. A system may be completely independent, its function uninfluenced by the state of any other neural system. This would be a truly modular system (see figure 1, upper panel). However, such a 'purely' modular system may be very uncommon. For example, measurement of psychophysical thresholds often reveal "interactions" between systems. This may be revealed by a slight improvement of performance of system-1 when system-2 is also active (consider stereo aiding a motion detection task, e.g. McKee et al, 1997), or it may be that the state of system-1 can influence or drive the state of system-2 (consider the cross-linked systems of accommodation and vergence). Thus, modules that are coupled to, or have informational links with, other modules may be more the norm (figure 1, lower panel).

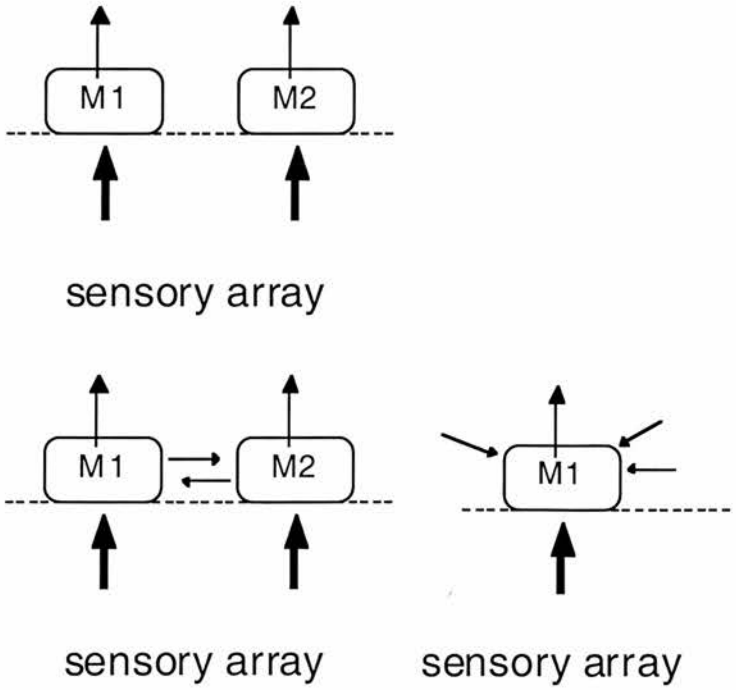


Figure 1. Information and systems. **Top,** two pure modular systems, M1 and M2. **Bottom left,** coupled-systems. **Bottom right,** a partially modular system

1.2.2 Activational autonomy?

Modules are informationally isolated from "central processes" such as memory. However, modules are not completely independent of central processes. Pöppel (1989; von Steinbüchel & Pöppel, 1993) proposed a set of "how" functions (see figure 2) that are responsible for (i) activation of the modular functions, and (ii) synchronisation of the modular functions. These functions are basic mechanisms underlying brain function, and are very different from Fodor's central processes such as memory. Pöppel's "how" functions are anatomically distributed. Therefore, brain damage may directly impair the functioning of specific functions (modules) but will also impair "how" functions, thus affecting modules indirectly.

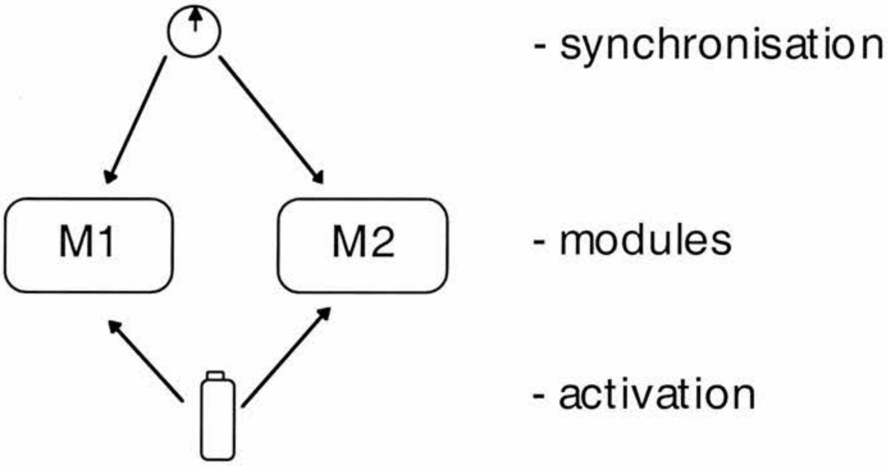


Figure 2. Activation and synchronisation of modules

1.2.3 Restricted access to information

Modules are "informationally encapsulated", they do not have access to all the information available in the brain. Fodor provides the vivid example of gently pressing the side of your eyeball. The world is seen to jump or distort. Although you know that you pressed you eyeball and this is the reason for the perturbation, you are unable to use this information to inhibit the percept of motion of the world. Fodor explains that such an independence of perceptual processes from central processes has two advantages. First, perception cannot be influenced by expectation or memory, therefore you see what is actually out there, not what you expect to see. This has obviously evolutionary advantages. Second, perceptual processing can be considerably faster if it does not need to involve referrals to central processes.

What about perceptual information? Do modules have access to all of the information available in the optic array? For instance would it make sense for a module for stereo depth to have access to colour information? At this point it is useful to bring in some ideas from another major theory of perception, Direct Perception. The tenets of Direct Perception are largely compatible with the theory of modularity. Cutting (1986) suggests that a modern formulation of direct perception (that due to Gibson) holds that perception relies on use of only a <u>subset</u> of the information that is actually available in the optic array (see Cutting, 1986, pp 242-252)[1]. Consider two popular examples of direct perception, the perception of direction of locomotion and the perception of time-to-contact. The literature can be read as indicating that the former relies on the instantaneous velocity flow field and makes no use of available information about ego-centric directions (see Warren, 1995 for a review of the literature on "heading"). The latter appears to rely on optical looming and makes no use of the available information from, for example, binocular disparity.

Although the attribution to Gibson, of the use of only a subset of the potential information, may be mistaken, there could be some sense behind the idea. Arguments can be advanced based upon speed and efficiency: Given two candidate systems that both return veridical information, if one system takes additional inputs then it will probably be more complex, less efficient and slower. This would suggest that systems which use the

---

[1] I am not clear whether Cutting (1986) interprets Gibson as not appreciating that a property may be specified by several sources of information in the optic array, or whether he interprets Gibson as appreciating this but advocating the theory that only a single source of information will be used. The former seems the least plausible so I assume the latter.

minimal set of information to perform the same function would be preferred. Further, the broader the range of information used by systems, the greater the degree of overlap of input and processing between systems. Overlap would be likely to inhibit modularity as Hebbian learning (the formation of neural links; Hebb, 1949) occurs when there are similar patterns of activation in two different areas. Development of links between systems undermines independence and hence modularity.

1.2.4 Breakdown of modularity?

Baron-Cohen et al (1993) suggest that synaesthesia is the breakdown of modularity. Synaesthesia is a percept in one modality triggered by stimulation of a different modality. For example, coloured-hearing is the experience of seeing colours when hearing a word or a sound. Synaesthetes report it as an involuntary percept and one that has occurred for as long as they can recall.

The existence of synaesthesia raises a whole tranche of questions about modularity. In addition to the breakdown of normal modularity interpretation, it has also been suggested that synaesthesia provides a window into the holistic nature of neural function (Cytowic, 1995). Segal (1997) offers the seemingly bizarre idea that a synaesthete possess an extra module. Baron-Cohen (1996) suggests that synaesthesia may be part of the development process, presumably the stage prior to individuation of modalities.

That modularity is so central to these explanations illustrates how compelling the theory of modularity has become. Unfortunately the extreme incompatibility of the interpretations of synaesthesia based upon modularity and the lack of obvious ways to discriminate between them is rather problematical.

Before attempting to resolve which of the above is most plausible it is worth taking a step back. A sceptic could argue that the existence of synaesthesia has yet to be proved. Synaesthesia may simply be the combination of very good memory and a vivid imagination. I considered it essential to convince myself of reliability of the synaesthesic experience. Together with Martin Corley I examined the stability of coloured-hearing synaesthesia (see 1.5.1). I concluded that synaesthesia is indeed a real phenomenon. My opinion about synaesthesia and modularity remains for now unformed. However, our results do prompt specific questions about the "wiring" of synaesthetes or the completeness of our understanding of the "wiring" of normals.

1.2.5 Plasticity of modularity?

Modularity is "innately specified" (Fodor, 1988). Is there any plasticity, is it possible to change the degree of modularity of a system? Brain damage suggests that the functional architecture is very rigid. When a specific function is lost, rehabilitation of function relies on "substitution", that is finding a different way of performing the same function. For instance, learning to lip-read if auditory function is impaired. It does not appear possible to train up other brain regions to take over the exact same function.

What about a more limited change in functional architecture? Is it possible to change, permanently or transiently, the independence or inter-dependence of systems? If two systems have cross-links between them, is it possible to temporarily turn the cross-links off so that the two systems can function independently?

A related question is: Is modularity prewired, or does it develop in response to the visual environment, do modules "embody" sensory invariants? If the latter, is it possible to change the course of development? By changing an infant's sensory environment could we change the resultant modularity of the system? There is evidence that the occipital lobe (the "visual cortex") may be able to take on some auditory function in the case of congenital blindness. This would indicate that the functional-anatomical structure of the brain is not completely prewired.

The development of new technology requires an answer to these questions which could have both practical and theoretical importance. Stereo head-mounted displays (HMDs) produce viewing demands that are in conflict with the functional organisation of the ocular-motor system. This is because HMDs disrupt the invariant relationship between focus and binocular disparity. When viewing an approaching object the eyes must change focus (accommodate) and swing inwards to place the object on the fovea of each eye (converge). There are neural cross-links between the two neural systems, such that a response in one system can drive a concomitant response in the other. When viewing a stereo display it is necessary to keep accommodation fixed on the plane of the display whilst varying vergence when changing regard from a close object to a far object. Such a situation would never arise in the natural world.

Is it possible to temporarily disable the cross-links? What would be the consequence of providing young children, with developing visual systems with such devices? Would the plastic, developing visual system be more ready to adaptation of neural links? Could normal visual development be permanently impaired? These are questions I have been concerned with. In chapter 2 there are brief details of the experimental work I have conducted, and there is also mention of current theoretical work employing neural networks (see 1.5.2).

1.2.6 Central "how" functions

Fodor's modules are independent from central processes, they are not influenced by expectation, and can't call on memory or problem solving resources. They are informationally independent. However, there is evidence that modules are not independent of all central functions. Modules are critically dependent upon central processes for activation and co-ordination. Pöppel (1994) provides a very vivid example of this dependence. Perception of vowel sounds relies on a fine temporal resolution of processing. If general temporal processing is impaired then this will impact language perception. Pöppel reports a case of agramatism (a problem with language processing) which was successfully treated by an intervention based solely on improvement of temporal acuity (Pöppel, 1994).

A recognition of the importance of these underpinning central processes motivated the research I was involved with into brain injury rehabilitation. We used an intervention that combined exercise and an enriched environment to target underlying deficits in activation and arousal that are associated with acquired brain injury. This research is summarised in 1.5.4. The outcome of the research supported the significance of central functions of activation and arousal.

1.2.7 Independence of modules from central processes

Unilateral visual neglect (UVN) is the fragmentation of the world into the neglected and non-neglected halves of space that may occur following acquired brain injury. Patients with UVN 'lose' space on the side contralateral to the major cerebral damage. UVN is normally classed as an impairment of a central process of representation or attention (Shallice, 1988).

Recall from the above, modules function autonomously, although they provide information to central processes their activational state should have no effect on a central process. An intriguing result was reported that raised questions about whether this was indeed so. An interaction between UVN and motion was reported (Mattingley, Bradshaw & Bradshaw, 1994). It appeared that the line dividing the neglected and the non-neglected could be changed by the activation of different modules. I investigated the reported interaction between UVN and motion. I employed a different measure and failed to find any such interaction. A concise summary of the issues and conclusions appears in the appendix (see 1.5.3).

1.2.8 Minimal information and modular processes

The potential advantage of an implementation of a function using the minimal necessary set of information was discussed above. The two examples of perception of locomotor direction and perception of time-to-contact were also mentioned. In the body of this thesis I examine both in some details.

A brief overview of perception of locomotor heading and perception of time-to-contact is given below. Each of the chapters that follow are derived from a paper and so include a brief explanation of the particular issue, and a concise summary and review of the pertinent previous research.

1.3 Perception of time to contact

How do we judge <u>when</u> something will hit us or when we will hit something? The "time-to-contact" (TTC) of an object is given by the distance of the object divided by the speed at which it is approaching.

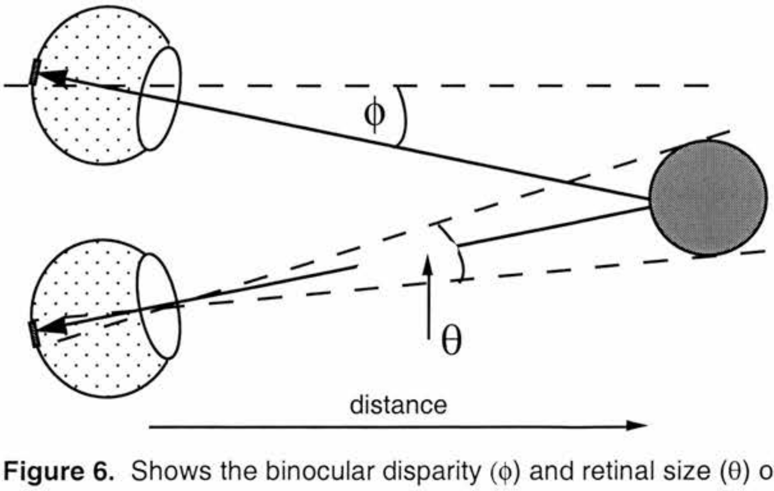


Figure 6. Shows the binocular disparity (ϕ) and retinal size (θ) of an approaching ball.

In the binocular array, optical size, θ, is approximately inversely proportional to the distance of the object, as is absolute optic disparity, ϕ (or binocular subtense). Therefore, TTC is specified by changing optical size (looming) and changing absolute optic disparity.

$$\text{TTC} = \text{distance/speed} = \phi / (d\phi/dt) = \theta / (d\theta/dt) \qquad [1]$$

1.5.1 The *tau* literature

Lee and colleagues, and others, have produced extensive research on the perception of TTC and the use of such information in the guidance of action. Infants, birds, children with Cerebral Palsy, runners, skiers, divers, punchers and jumpers, have all been studied (for example see Lee, 1976; Lee, Lishman & Thompson, 1982; Lee & Reddish, 1981; Lee, Young, Reddish, Lough & Clayton, 1983). Additionally, Tresilian has written in copious detail on the theoretical issues associated with the perception of TTC (see for example, Tresilian, 1990; Tresilian, 1995).

1.5.2 Regan and colleagues

Regan and colleagues have thoroughly studied perception of motion-in-depth, thresholds, mechanisms, the combination of looming and changing disparity (see for instance Regan & Beverley, 1979; Gray & Regan, 1997).

Surprisingly, there is very little overlap between the two bodies of literature. From the work of Regan it seems clear that disparity will be used in perception of TTC. However *tau* studies appear to ignore this source of information.

It is not clear why this is so unless it is motivated by the hypothesis set out above that humans only exploit a subset of the information available to them.

What of disparity and the evidence for the exclusive use of looming? In a recent review Wann concluded that the evidence in support of the exclusive use of looming was weak (Wann, 1996). There are several studies that point to a role for disparity, Heuer (1992) manipulated both looming and disparity and concluded that both are used in the perception of TTC. Judge & Bradford (1988) had participants wear telestereoscopes, which change disparity. They found that telestereoscopes perturbed catching of a ball, thus indicating that disparity has an influence in the task of ball catching. More recently, Gray & Regan (1997) have documented the influence of both disparity and size on perceptual judgements of TTC.

1.5.3 Timing of interception away from the face

Most models of TTC are concerned with perception of TTC with the face. Using TTC_{Face} to time interception of a ball in front of the face would lead to errors. The error will be a function of the distance of the interception point from the face and the speed of the approaching object.

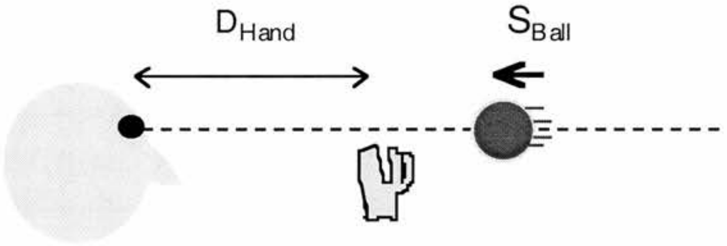


Figure 7: A ball approaching with a speed S_{Ball} and the hand at a distance D_{Hand} from the face.

$$TTC_{Hand} = TTC_{Face} - D_{Hand} / S_{Ball} \qquad [2]$$

Where TTC_{Hand} is the TTC with the hand, D_{Hand} is the distance of the hand from the face, and S_{Ball} is the speed the ball is travelling towards the face.

How does an observer actually calculate $\mathrm{TTC}_{\mathrm{Hand}}$? It could be done via the route described in equation [2], by correcting an estimate of $\mathrm{TTC}_{\mathrm{Face}}$. Alternatively it may be done directly.

Again, conspicuously most papers avoid this problem. A few suggestions (e.g. Bootsma & Oudejans, 1993; Peper, Bootsma, Mestre & Bakker, 1994) have been offered to the perception of $\mathrm{TTC}_{\mathrm{Hand}}$, but most are special cases, such as when the ball approaches from an eccentric position and arrives at the hand (in which case it is possible to use lateral gap closure to estimate TTC).

There is a simple solution when the hand is visible in front of the face. Here an observer can use relative disparity (the difference in disparity between the hand and the ball). When the hand is not visible, then it becomes very difficult as there is no reference for optical specification and the observer must use *intended* hand position in timing.

In later chapters (chapters 3 and 4) I explore some of the issues raised above.

1.4 Perception of locomotor direction

When walking across a field how do we guide ourselves across to the gate and not end up in the hedge?

The answer which has become commonly accepted is that we make use of the changing pattern of light on the retina. Movement forward produces a radial pattern of motion or 'flow'. A step sideways streaks laminar motion across the retina. Walking a curved trajectory also produces a characteristic pattern.

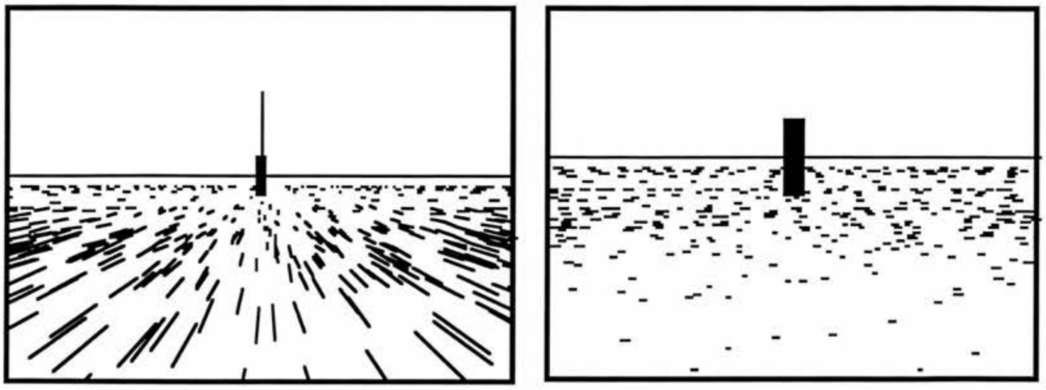


Figure 8: Left, a step forward and **right**, a step sideways

In the late 40's and 50's Grindley (cited in Mollon, 1997), Gibson (1950) and Calvert (1950) noted that certain patterns or invariants in the distribution of motion indicated the direction of locomotion (notably the 'focus of expansion, the point from which all motion appears to radiate, see fig 8). They hypothesised that humans may use these patterns to judge where they are going.

The first real experimental research into this hypothesis was conducted in the 60's. Carel (1961, cited in Cutting 1986) examined judgement of locomotor direction during simulated approach to a surface. Carel found an accuracy of the order of about 4 degs. A number of studies that followed Carel's over the following years, most also reported poor performance (Johnston et al, 1973; Llewellyn, 1971; Warren, 1976; Regan & Beverley, 1982).

These results were troublesome because they correspond to performance levels below that required for walking, running or driving (Cutting, 1986, estimates that an accuracy of 1-3deg is necessary).

Further problems were apparent when Regan & Beverley (1982) pointed out that use of the flow field was not as simple as originally conceived. Their point was that when an eye-movement is made it complicates the flow-field and determining the direction of travel from the gradient of motion is no longer possible (see fig 9 below).

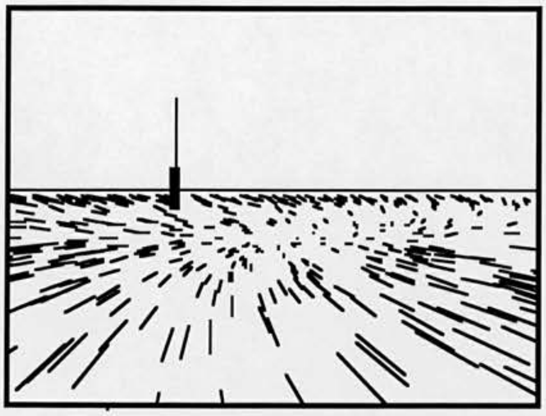


Figure 9: Moving towards the tower (solid black object) whilst maintaining fixation on a point infront of it, on the ground. Note the focus of expansion no longer indicates the direction of travel.

It was demonstrated to be mathematically possible to 'decompose' the complicated observer movement + eye-movement field to remove the confounding eye-movement (Longuet-Higgins & Prazdny, 1980). However, this appeared to undermine the simple flow hypothesis.

Fortunately, Warren & Hannon (1988) rode in to save the shaky looking flow model. They demonstrated first that observers could judge their direction of travel with the necessary high accuracy when presented with displays simulating motion across a ground plane. Second, they showed that observers could do this even in the presence of an eye-movement. Warren & Hannon added a 'simulated eye movement', this means that they added a rotational component to the flow field that corresponded to that which would be added if the observer made an eye-movement. Warren & Hannon added a simulated eye-movement which produced the complicated flow field without getting the observers to actually move their eyes. They demonstrated that observers could judge their direction of locomotion purely on the basis of optic flow and without having to resort to use of extra-retinal eye-movement information (either a copy of the signal to move the eye and/or a signal indicating the change in position of the eye).

Warren's work sparked a large amount of further research (on psychophysics, computational models and neural substrates). Several issues were followed through, the major ones being the role of extra-retinal information and the role of depth.

Royden, Banks & Crowell (1992) suggested that Warren & Hannon's finding did not hold at higher rotation rates typical of natural eye-movements and that at high gaze rotation rates it was necessary to use eye-movement information. van den Berg (1993) cut across this argument by pointing to a potentially confounding variable, the 3D structure of the scene through which simulated movement occurred. van den Berg noted that a ground plane provides information about the distance of each point in the scene. When information about the depth of the constituent elements is present he argued, high performance is still possible without eye-movement information. This line of reasoning was extended by several other papers which showed that the addition of disparity-defined depth appeared to improve performance (van den Berg & Brenner, 1994).

The argument about eye-movement information now appears to be resolved with most researchers believing that it at least aids perception of locomotion direction, even if they continue to dispute whether it is strictly necessary.

The role of depth information is the subject of Chapter 5 so I will not dwell upon it longer here.

Whilst the flow debate was occurring, Cutting (1986, 1991) was championing an alternative theory. He suggested that rather than recover direction of locomotion from a complicated flow field, humans may instead use the relative motion of objects in the scene to determine their direction of travel. Cutting pointed out that 'differential motion parallax' (DMP), that is, the relative displacements of objects within a scene, is unaffected by an eye-movement. Cutting has argued his case almost on his own, with most researchers opting to ignore DMP and concentrate on optic flow.

There has also been related research concerned with car driving (rather than locomotion on foot) and the use of "tangent points" (e.g. Land & Lee, 1994) and splay (the angle of the sides of the road) (e.g. Beall & Loomis, 1996) when driving along a road.

Most of the research that has been conducted on locomotor direction has relied on perceptual judgement tasks. This was probably in some part due to technical limitations. Materials for perceptual judgements are essentially just brief film clips. These can be generated in advance over a period of time and then played back to the observer at a high-frame rate (much as a cartoon is turned from stills into a "movie").

However, there are some problems associated with trying to extrapolate from results obtained using perceptual judgements to active perceptual control. Although judgements may provide insight into what a human can do, they do not indicate what a human actually does in a natural context.

There are a number of potential strategies available to the human actively controlling locomotion that are not available to the observer making judgements. Most critical point is that during active control, feedback is available. Therefore, a source of information that only provides crude "going left" vs. "going right" information can be used as the observer can progressively refine their response. During a perceptual judgement such information would be useless and it is necessary to use a source of information that provides "ratio-level" information - tells you exactly where you are going. Therefore, the very different requirements of the two tasks could lead to use of very different strategies or mechanisms in the two cases. Wann, Rushton & Lee (1995) report a study that may indicate that different mechanisms are employed during active control and perceptual judgement. They found that, in contrast to a result reported by Warren et al (1991) on perceptual judgements, the instantaneous velocity field does not provide the optimal input for active control of locomotor direction.

For the reasons noted above, the studies described within the following chapters employ active control rather than perceptual judgement. Chapter 4 examines the role of depth in the perceptual control of locomotor direction. In Chapter 5 prism glasses are used to explore the role of ego-centric direction of the perceptual control of locomotor direction on foot (i.e. when walking). Chapter 6 provides a more thorough description of a model of control locomotion on foot based upon ego-centric direction.

1.5 Extended Abstracts

This section contains four extended abstracts describing research that I have been engaged in that has a bearing on the modular nature of the visual system.

1.5.1 Synaesthesia

'[Synaesthesia] denotes the rare capacity to hear colours, taste shapes or experience other equally startling sensory blendings' Cytowic (1995). Baron-Cohen suggests that synaesthesia is the 'breakdown of modularity'.

Synaesthesia is very rare (1:2000; Baron-Cohen et al, 1996), 'coloured-hearing', the involuntary percept of colour upon hearing a sound or word is by far the most common form. Baron-Cohen and colleagues (1987, 1993, 1996) have reported a number of cases of coloured-hearing synaesthetes who, given a list of 100+ words or letters, can describe the colour they see (for instance "FEAR is a mottled grey"). When retested several weeks later with the same set of items they produce matches that show a remarkable similarity to those from the first session.

Although this performance is impressive, a sceptic can suggest that synaesthetes do not really see colours in response to letters but rather just report that they do. Some people manage to recall pages of telephone numbers. It may be that the synaesthete simply consigns to memory the 100 or so colour descriptions that they generated during the first session.

An interesting brain imaging study has recently been reported on synaesthesia. Paulesu et al (1995) conducted brain imaging during coloured-hearing synaesthesic experiences. Changes in cerebal blood flow were found in associative visual areas. However, no activation was noted in the primary visual areas, including no significant increase in activation in V4, the putative colour area (Zeki, 1993). This is intriguing - how does a vivid percept of colour arise without the involvement of what we take to be the neural areas responsible for vision and the perception of colour?

One interpretation of the imaging data could be that it is compatible with the use of a non-synaesthesic strategy based upon memory.

Together with Martin Corley, I investigated the stability of coloured-hearing (Rushton & Corley, 1998a, 1998b). Rather than use a colour-naming task we planned to have synaesthetes match to chips in a Munsell Colour Atlas. We started with the premise that if synaesthetes just recalled verbal descriptions and then visualised the colours, then this would probably be revealed in several ways: First, the colours recalled would probably be prototypical hues (red, green, yellow etc.). Second, synaesthetes would be unlikely to be able to visualise exactly the same colour each time (for example "red", even "blood red" is not a very precise colour).

We tested three self-nominated synaesthetes using a Munsell colour atlas with 1150 colour chips, a list of 20-30 items (letters, numbers, days of the week). The synaesthetes were asked to locate the best match for each item in the Munsell atlas (see fig 1). We then retested with the same set of items a week later. The assumption was that if our subjects were using a memory strategy that the matches would be approximately the same. If true synaesthesia was occurring then the matches would be considerably better.

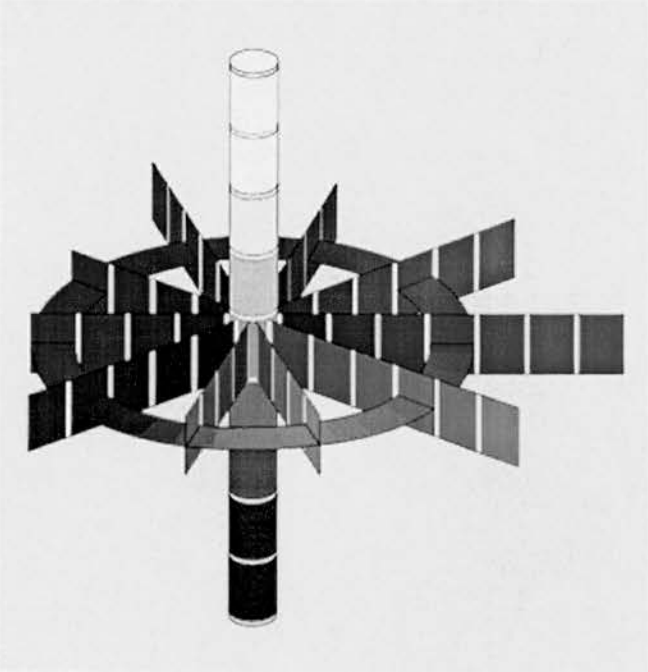


Figure 1: A visual representation of the three dimensions of Munsell colour space. Hue is around the pole, chrome is outwards from the pole and value is up and down the pole.

Synaesthetes responses were translated from Munsell colour space to different colour space, CIELAB. This is because Munsell is not a perceptually uniform space, CIELAB is somewhat better. For each item we calculated an error based on the Pythagorean distance in colour space between the matches from the two sessions. The distribution of colours and errors was closely examined. The summary results for three synaesthetes, L.F., L.B. and C.C. are shown below in table 1.

Two colours that are 10 units apart are considered by most observers, under most viewing circumstances as the same colour. Two chips need to be at least 40 units apart in colour space to be reliably seen as different (van Laar, *personal communication*). It can be seem from table 1 below that synaesthetes performance is very impressive. Closer examination of the errors showed only one synaesthete, once, picked a colour chip 40 units away from the chip chosen in the first session.

	mean	SD	N
L.F.	12.9	11.3	24
L.B.	16.7	12.7	22
C.C.	5.5	6.8	21

Table 1. Mean and standard deviations of 3D distances (ΔE) in CIELAB colour space between colour matches in first and second session.

These results are considerably better than would be expected from a verbal description-visualisation strategy.

It is possible to cheat with the Munsell colour atlas by remembering the spatial location of the chip or its co-ordinates. Use of such a strategy should be revealed by examining the distribution of errors. A very careful examination of the distribution of errors produced no indication of use of memory-based strategies.

We were forced to conclude that synaesthesia is a real phenomenon. The result is perplexing. From the empirical data it appears that synaesthetes do have a vivid, concrete percept of colour. However, the imaging studies show no activation of primary visual cortex. Given that the role of V4 in colour perception is so widely accepted (Walsh, 1996) and indeed that it has been suggested, that conscious experience requires activity in primary visual area V1 (Crick, 1994), this prompts several awkward questions: Are synaesthetes 'wired up' differently to the rest of us? Is our knowledge of neural substrates for colour perception incomplete? Unfortunately, although our data establishes the validity of the questions, it does not provide the answers.

1.5.2 Accommodation and vergence

1.5.2.1 Stereoscopes

When Wheatstone built his first stereoscope, he created a device that disrupted a natural invariant relationship between binocular disparity and focal distance. Normally, as an object moves in from a distance an observer is required to accommodate (change the focal power of the lens in their eye) and converge (swing their eyes inwards so as to keep the object of interest on the fovea of each eye). Accommodation and vergence thus co-vary (see fig 2).

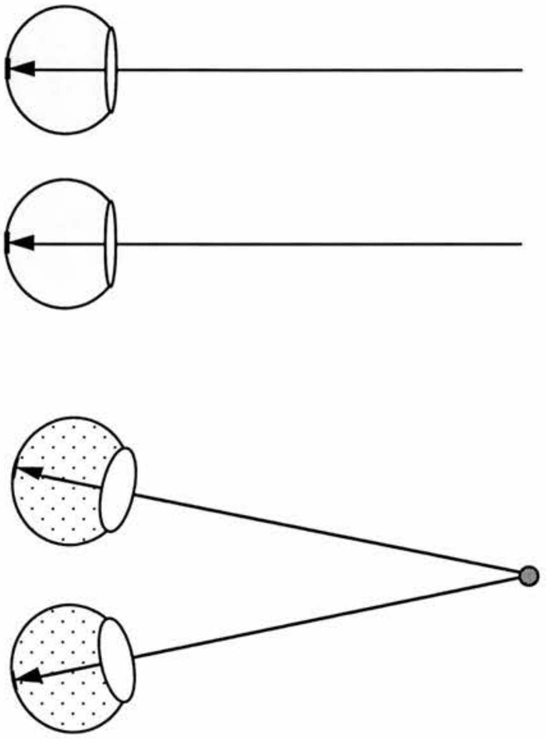


Figure 2: looking at a far (top) and near (bottom) object. Note that accommodation (focus) and vergence co-vary.

The visual system embodies this invariant relationship by "coupling" the systems with cross-links between them.

When viewing images using a stereoscope, the eyes have to verge whilst accommodation remains fixed on the plane of the screens (fig 3).

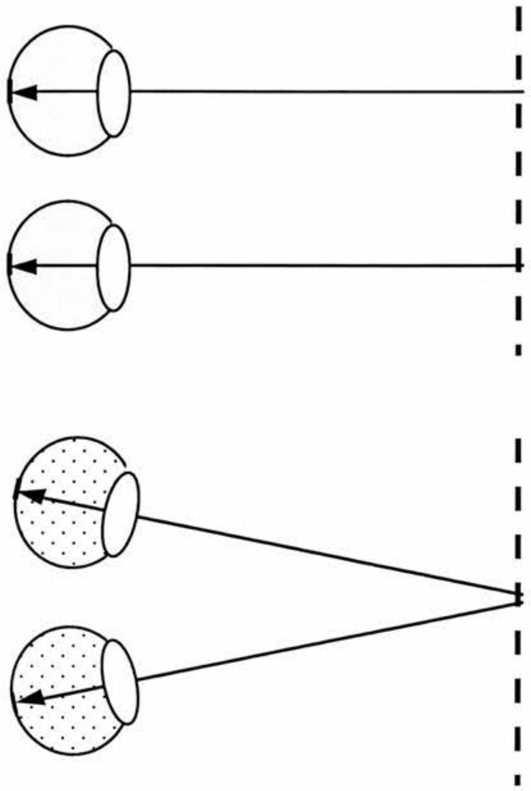


Figure 3: looking at a far (top) and near (bottom) object using a stereoscope. Note that accommodation (focus) is required to remain fixed whilst vergence changes.

This should be difficult because of the "coupling" between the systems. Shallice (1988) notes that one feature of coupled systems is that it is not possible for them to take up contradictory states. Observer's do manage to view images in stereoscopes, presumably due to some slack in the systems (dead-zones like depth of focus) and slack in the couplings (this is known clinically as "fusional reserves" or "relative vergence and accommodation" and indicates how much the two systems can be pulled out of sync before blur and then double vision occur).

Some people find fusing stereo pairs easier than others and practice generally facilitates fusion. Clinical data shows that there is a large amount of variability in individuals ability to dissociate accommodation and vergence. Orthoptics (vision training) is also based upon the assumption that it is possible to improve this ability. It would seem likely that variation in ability to dissociate accommodation and vergence and the variation in ability to fuse stereo pairs is related to the strength of the coupling between the two systems.

1.5.2.2 Head-Mounted Displays (HMDs) and Virtual Reality

Stereo head-mounted displays rely on the same principle as the stereoscope but are in rather broader usage than stereoscopes. Users complain of eye-strain and demonstrate measurable changes in binocular function (Mon-Williams, Wann & Rushton, 1993; Rushton, Mon-Williams & Wann, 1994; Wann, Rushton & Mon-William, 1995; Mon-Williams, Wann & Rushton, in-press; Rushton & Riddell, in-press). It seems likely that this may be in part due to the challenging visual stimulus (there are also often problems with display quality).

1.5.2.3 Individual variations, coupled systems and adaptation

Comfort and safety in the use of HMD's is obviously a problem in itself. However, it is also an interesting theoretical matter. There is considerable variation amongst individuals in symptoms and ophthalmic changes after use of an HMD. There is an indication that repeated exposures may reduce symptoms. What accounts for individual variations and what is the mechanism that reduces problems with repeated exposure? Is the former determined by the strength of the coupling between the systems? Does the latter result from the development of capacity? Does the visual system develop a dual-adaptation facility? If so how does it work? The degree of modularity is supposed to be fixed, does coupling get turned on and off with a dual adaptation? What would happen to children who have developing and more plastic visual systems than adults? Would they have less problems because of the greater plasticity, or could it disrupt and change the development of normal binocular vision?

Ethical problems obviously beset the empirical testing of such questions. Recently I have started to investigate these matters by an indirect route. With Patricia Riddell and John Bullinaria, I have been exploring neural network modelling of the development of accommodation and vergence. The hope is that some of the questions about independence of systems, mechanisms of adaptation, and influence of developmental plasticity may be answered by training and challenging neural networks, rather than real human visual systems (Rushton & Riddell, in-press; Riddell, Bullinaria & Rushton, in-press; EPSRC Grant GR/L82274, Riddell & Rushton, 1998).

1.5.3 Unilateral Visual Neglect

Unilateral visual neglect (UVN) is the fragmentation of the world into the neglected and non-neglected halves of space following acquired brain injury.

1.5.3.1 UVN background

Immediately following an acute hemispheric stroke over 60% of patients exhibit unilateral visual neglect (Stone, Halligan & Greenwood, 1993) - they disregard objects in the space contra-lateral (opposite) to the major cerebral damage. They literally appear to lose half of their world. UVN is not a visual field deficit. Although often accompanied by hemianopia (the loss of half the visual field), unilateral visual neglect dissociates from it. Classic, textbook signs of neglect are eating food on just one side of the plate and then believing it to be empty, combing the hair on just one side of the head. Clinical measures such as line bisection (the centre of the line is marked by the patient and found by the examiner to be far off to the ipsi-lateral side) and object cancellation (objects on the side contra-lateral are omitted when instructions are given to cross-out all the objects present, e.g. all the 'o's on a sheet of paper containing 'o's and 'x's) demonstrate equivalent behaviour (see Appendix, UVN: Line bisection and cancellation tasks). In the figures below, the patient has left neglect.

In the first figure (figure 4) he has taken a photograph with the subject "centred". He has not noticed the area of space to his left and believes the subject to be in the middle of the viewfinder.

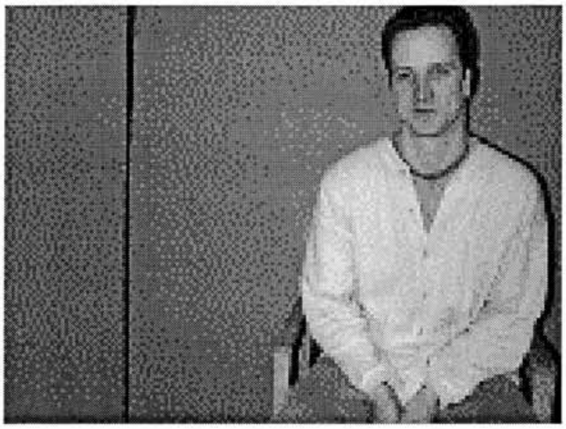


Figure 4. A photograph taken by a patient with left UVN. Instructions were given to centre the subject in the view finder. It can be seen that the subject is located well to the right of the mid-line. (Rushton, Johnson & Wann, 1996)

In the second figure (figure 5) he is trying to pick up the bar from the centre but misses the left part of the bar and so reaches away from the centre.

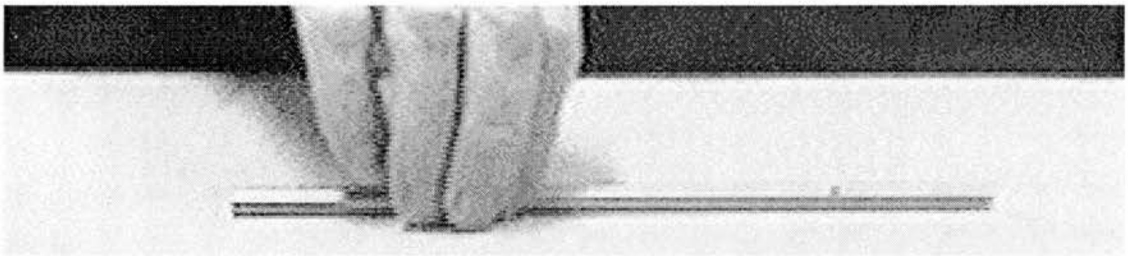


Figure 5. A patient with left UVN attempting to pick up a bar between finger and thumb from the "centre". (Rushton, Johnson & Wann, 1996)

Traditional accounts would have it that the left side of the bar is 'neglected' or not seen, thus the patient is reaching for the middle of the bar that they perceive. An alternative theory (Milner, Harvey, Roberts & Forster, 1993) suggests that the world becomes 'compressed' so for example in the case of line bisection, the true ends of the bar are perceived but the left half appears shorter than the right.

1.5.3.2 Neglect across visual modalities?

Unilateral visual neglect is classed as an attentional or representational deficit. Therefore, UVN should be unaffected by the activation of differing perceptual modules. For instance a line specified by a luminance change and a similar line defined by a colour change should both be (mis-) bisected at the same place. However, a few intriguing results were reported in the literature about UVN and motion that suggested that this assumption may be wrong, that the 'fracture line' between the neglected and non-neglect fields may vary dependent upon the visual modalities involved.

1.5.3.3 Interaction with motion

Mattingley, Bradshaw & Bradshaw (1994) presented patients with a line bisection task on a computer. The background comprised of small dots above and below the line that were either static, leftward moving or rightward moving (and a control, neutral background). They found that patients were affected by global motion in the background; a leftward moving background showing a significant shift in bisection error. A very similar study by Pizzamiglio, Frasca, Guariglia, Incoccia & Antonucci (1990) looked at the effects of "optokinetic stimulation". Patients were asked to bisect a luminous strip in half. The strip was surrounded by luminous dots that were either fixed, left or right moving. It was found that neglect was reduced with a leftward moving background.

There are four simple explanations for this interaction between neglect and visual motion: (i) a simple change in bias due to motion dragging the eye (optokinetic nystagmus) or attention in the direction of dot motion; (ii) the invocation of different perceptual-attentional systems because of the addition of a movement attribute to the test element (relative motion against the background flow makes the line visible to the motion system; see Appendix, UVN: Motion as an independent process); (iii) reduction in attentional bias between the neglected and non-neglected fields due to visual motion acting as common stimulus to attention across the whole test display; (iv) finally, visual motion may shift the perceived mid-line back towards the real mid-line by either feed into a common pathway to the vestibular signal (Wertheim, 1994) or changing the 'straight-ahead' position of the eyes either of which would change the relationship between body frames of reference.

1.5.3.4 What I did and found

Previous research used a line bisection task. I used a cancellation task (see Appendix, UVN: Line bisection and cancellation tasks) instead, in the hope that it would clarify whether super-imposed motion produced a bias in response or actually a reduction in neglect. With a cancellation task it is possible to distinguish between the two: with a bias the same percentage of elements is cancelled by the experimental manipulation effects, with a reduction in neglect the total number of elements cancelled increases.

I was unable to collect any data that suggested that there was a reduction in neglect or even a change in bias. I also had difficulty replicating previous line bisection results due to the problems of fluctuations (see Appendix: UVN fluctuations).

It is not possible to conclude whether there is an interaction between motion and UVN on the basis of my studies (Rushton & Wann, MRC 9520612). Negative conclusions are always very hard to substantiate as it only takes one patient to demonstrate a dissociation or interaction, but it takes a very large number of patients to suggest a dissociation or interaction may not be possible. All that I can note is that I did not come across any cases in which there was a substantial difference between performance with and with-out motion.

1.5.4 Brain injury rehabilitation

Whilst pursuing my research into unilateral neglect I got drawn into research on general brain injury and rehabilitation. My clinical collaborator, David Johnson, together with his academic collaborator, David Rose have been developing an idea that following brain injury, "activation-arousal" is depressed and that this underlies the failure of rehabilitation (e.g. Rose & Johnson, 1994). Fatigue, lack of concentration and depression characterise all cases of brain injury. These undermine the application of standard rehabilitative interventions. If a patient is to benefit from interventions it is necessary to tackle these underlying problems by targeting the neurophysiological changes and under-lying problem of "activation-arousal". von Steinbüchel and Pöppel (1993) examined approaches to rehabilitation and proposed a very similar idea to that developed by Johnson and Rose.

In terms of modularity we can say that unless you work on the underlying cognitive resources you are not going to have very much success working on modular functions. Further, you may produce improvements in modular functions by concentrating solely on the general cognitive resources (Pöppel , 1994).

So how do we turn this idea into a rehabilitative intervention? Animal research, studies in sports journal and every day experience all indicate that physical exercise is beneficial. It has been reported to improve cognitive function and mood as well as build physical stamina and strength. Animal research shows underlying neurochemical and neurostructural changes associated with physical exercise (e.g. Neeper et al, 1995).

"Enriched environments" are also associated with benefits. Again animal studies demonstrate superior ability on maze tasks and also changes in neurophysiology associated with enriched environments (e.g. Renner & Rosenzweig, 1987). Patients spend most of their time stationary in the ward, passively passing the time of day, apart from during visits by family and friends or physio-therapy, occupational therapy or neuropsychology sessions.

We built a computer generated Virtual Environment through which a patient could move and provided a number of simple tasks to test cognitive abilities and motor control. The patients "rode" through the VE which was displayed on a large monitor in front of them by cycling on an exercise bike (see fig ; Johnson, Rushton & Shaw, 1996a, 1996b).

Figure 6. Instrumented bike and large screen display.

This allowed us to combine physical exercise and a stimulating environment. We ran some initial pilot studies which looked very promising (Johnson, Grealy, Rushton & Soryal, 1997). Funding was obtained for a commercial "VR bike" and a small scale study was conducted by Madeleine Coleman. Significant improvements were found on a number of neuropsychological measures (Coleman, Johnson, Soryal & Rushton, 1998; Coleman, Johnson & Rushton, 1998)

The results were impressive, but a larger scale, properly balanced trial is required to determine whether the result is robust. Pöppel (1994) has already reported the effect of training on temporal acuity (a general cognitive function) on a case of aggrammatism. Robertson and colleagues (e.g. Robertson, Tegner, Tham, Lo & Nimmo-smith, 1995) have demonstrated the relationship between unilateral visual neglect (apparently a specific deficit) and general attention. These results highlight the relationship between "independent" modules and underlying cognitive functions.

1.5.5 Appendix

A1.1.4 A synaesthete's (L.F.) descriptions of the visual experience associated with letters of the alphabet

These descriptions were sent to the investigators several weeks after the experiments. The letters are those that L.F. chose to describe (spelling as per original).

a - solid, yellow a shape, smooth rounded edges
f - green/grey, texture like a mountainside of fir trees seen from above
e - watery indescent blue
j - pale purple, not very influential, thin and reedy
k - very strong metallic dark gold/brown hard edges
n - dark grey/black
p - purple/pink air, no solid structure. warm, floating soft woolly lumps
s - white, quite cold and hard
u - optimistic yellow
v - grey, powdery like tailors chalk but a bit more dusty
x - middling grey
z - very reflective metal, very sharp edges

Most of the letters are seen just as coloured letters with differing degrees of hardness/softness and warmth/cold. The rigidity of the edge also differs with some (e.g. z) having razor sharp edges while others (e.g. p) hardly has edges at all.

A1.2.2 UVN: Line bisection and cancellation tasks

Line bisection is one of the classic tests of UVN. A patient is asked to mark the centre of a line presented to them. They will tend to mark a point away from the true centre. In the case of left UVN (following right hemisphere damage), the patient will mark towards the right of the bar, seemingly 'neglecting' the left side. The bisection point does depend upon several factors including the length of the line (see Halligan, 1995 for a comprehensive review).

A cancellation task presents a regular or irregularly spaced grid of elements. The elements may be lines, letters or shapes. The task for the patient is to mark or strike out all of the target elements. The target elements may all be targets or only a subset. Patients with neglect commonly make omissions on the side contralateral to the lesion, so typically a patient with a right hemisphere lesion will miss targets on the left of the page.

A1.2.4 UVN: Motion as an independent process

Evidence exists for the neuro-anatomical separation of processing of motion from colour and form (see Zeki & Shipp, 1988 for a review). Its status as a fundamental visual dimension and its functional separation is demonstrated by phenomenon such as motion after-effects (see Nakayama, 1986 for a summary). Motion also appears to have a privileged status in attention and visual search as is demonstrated by studies such as McLeod, Driver & Crisp (1988).

Riddoch (1917) first noted that following cerebral insult, movement may be recognised as a special visual perception that maybe retained despite loss of perception of static objects and scenes. He also noted that *"appreciation of movement returns before the object as such is recognised, if recovery of vision is occurring"*. Blythe et al (1987) examined patients with homonymous visual field scotomas following retrogeniculate lesions and found that 5 out of 25 were sensitive to bright flashes and simple motion in the blind field. Mestre et al (1992) demonstrated that a cortically blind patient, with negligible static spatial perception, retained perception of speed and direction from a complex (dot) flow field similar to that produced by locomotion. Ceccaldi et al (1992) report a patient with only 2 deg of macular vision and perifoveal sparing between 10 to 30 deg in the left inferior quadrant who could perform visually guided motion and consciously perceive motion of objects in blind parts of visual field. Conversely, Zihl, von Cramon & Mai(1983) document the case of a patient with posterior brain damage for whom perception of motion was selectively impaired.

A1.2.5 UVN: fluctuations

Only two papers out of the hundreds published make (passing) note of the fluctuation of UVN. (They suggest that the variance associated with UVN is due to a reduced sensitivity and so consequently larger Weber fractions.) This is very strange and I believe unfortunate as I now believe the fluctuations associated with UVN are potentially one of its most important features.

The downside of fluctuations

If a UVN patient is asked to repeatedly bisect a line then he or she will show a very large variance on response. Some of this variance can be accounted for by differences in line length and use of strategies. However most variance goes unexplained and very often unreported. It appears fairly standard practice to not report a measure of distribution in papers on UVN, only means or variance of a group. Most of the results reported in the literature either lack statistical support or report statistical significance on the basis of group results. There are very few interventions or manipulations that produce 'clinically significant' changes in response. Clinical significance is defined in many ways but the most common is that a change in performance of a patient is clinically significant when the patients response is now 2 standard deviations beyond the mean of the population from which the patient was drawn. With UVN it is near impossible to find such a change as the distributions are so broad.

The upside of fluctuations

Examination of responses often shows a repeating pattern. I have not quantified this with auto-correlations but the repetitions can be visible by eye when plotted against time.

It may be that fluctuations are characteristic of severe damage and prove useful as a prognostic tool. Alternatively, it is known that a patient's awareness of UVN is a major predictor of outcome. Fluctuations in response may be because of fleeting awareness of parts of neglected space or reflect the iterative use of strategies by a patient to try to overcome the UVN. In this case fluctuations may be an indicator of a positive outcome. Lastly from a different perspective, fluctuation is characteristic of a dynamical system prior to a bifurcation point. Therefore an increase in fluctuation may indicate that there is likely to be a qualitative change in the patient's state. The onset or reduction of fluctuation may indicate periods when rehabilitative intervention may be most useful or should stop. It would be necessary to conduct some tracking studies on a number of UVN patients to try to distinguish between these alternatives.

1.6 References

Baron-Cohen, S. (1996). Is there a Normal Phase of Synaesthesia in Development?, *PSYCHE*, 2(27).

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J. & Bolton, P. (1996). Synaesthesia: prevalence and familiality. *Perception*, 25, 1073-1079.

Baron-Cohen, S. & Harrison, J.E (eds) (1997). *Synaesthesia: classic and contemporary readings.* Blackwell Publishers, Oxford, UK.

Baron-Cohen, S., Harrison, J., Goldstein, L.H. & Wyke, M. (1993). Coloured speech perception: is synaesthesia what happens when modularity breaks down? *Perception*, 22, 419-426.

Baron-Cohen, S., Wyke, M.A. & Binnie, C. (1987) Hearing words and seeing colours: an investigation of a case of synaesthesia. *Perception*, 16, 761-767.

Beall, A.C. & Loomis, J.M. (1996). Visual control of steering without course information. *Perception*, 25, 481-494.

van den Berg AV. (1993). Perception of Heading. *Nature*, 365: 497-498.

Brodatz, P. (1966). *Textures: A Photographic Album for Artists and Designers.* Dover Publications, New York.

Bruce V, Green PR, Georgeson MA. (1996). *Visual Perception: Physiology, Psychology & Ecology.* Psychology Press, Hove, UK

Butter, C.M. & Kirsch, N. (1995). Effect of Lateralised Kinetic Visual Cues on Visual Search in Patients with Unilateral Spatial Neglect. *Journal of Clinical and Experimental Neuropsychology*, 17, 856-867.

Calvert E.S. (1950). Visual aids for landing in bad visibility, with particular reference to the transition from instrument to visual flight. *Transactions of the Illuminating Engineering Society, London* , 15, 183-219.

Crick, F. (1994). *The astonishing hypothesis: The scientific search for the soul.* Charles Scribner's Sons, New York.

Cutting, J.E. (1986). *Perception with an eye for motion.* MIT Press, Cambridge, Massachusetts, USA.

Cutting JE, Springer K, Baren PA & Johnson SH. (1991). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: Human Perception & Performance*, 121, 41-72.

Cytowic, R.E. (1982a). Synesthesia: I. A review of Major Theories and Their Brain Basis. *Brain and Cognition.* 1, 23-35.

Cytowic, R.E. (1982b). Synesthesia: Psychophysical Relationships in the Synesthesia of Geometrically Shaped Taste and Colored Hearing. *Brain and Cognition.* 1, 36-49.

Cytowic, R.E. (1989). *Synesthesia: a union of the senses.* Springer-Verlag, New York.

Cytowic, R.E. (1993). *The Man Who Tasted Shapes.* Abacus, London.

Cytowic, R.E. (1995). Synesthesia: Phenomenology and neuropsychology. *PSYCHE*, 2(10). Also reprinted in Baron-Cohen, S. & Harrison, J.E (eds) (1997). *Synaesthesia: classic and contemporary readings.* Blackwell Publishers, Oxford, UK.

Fodor, J.A. (1983). *The Modularity of Mind.* Cambridge, Massachusetts.

Frith, C.D. & Paulesu, E. (1997). The physiological basis of synaesthesia. IN In Baron-Cohen, S. & Harrison, J.E (eds). *Synaesthesia: classic and contemporary readings.* Blackwell Publishers, Oxford, UK.

Gibson J.J. (1950). *The perception of the visual world.* Houghton Mifflin, Boston.

Goodale, M.A. & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in* Neuroscience., 15, 20-25

Johnston, I.R., White, G.R. & Cumming, R.W. (1973). The role of optical expansion patterns in locomotor control. *American Journal of Psychology*, 86, 311-324.

Gray, J.A., Williams, S.C.R., Nunn, J. & Baron-Cohen, S. (1997). Possible Implications of Synaesthesia for the Hard Questions of Conciousness. IN In Baron-Cohen, S. & Harrison, J.E (eds). *Synaesthesia: classic and contemporary readings.* Blackwell Publishers, Oxford, UK.

Grealy, M.A., Johnson, D.A. & Rushton, S.K. (under-review). Improving Cognitive Function Following Brain Injury: The Use Of Exercise And Virtual Reality. *Archives of Physical Medicine and Rehabilitation*

Grealy, M.A., Johnson, D.A., Soryal, I. & Rushton, S.K. (1998). The effects of exercising in virtual environments on cognitive rehabilitation following traumatic brain injury. *Proceedings of the BPS Conference,* March, 1998.

Hebb, D.O. (1949). *The organisation of behaviour.* Wiley, New York.

Heuer, H. (1993). Estimates of time to contact based on changing size and changing target vergence. *Perception* 22, 549.

Johnson, D.A., Rushton, S.K. & S haw, J. (1996). Virtual Reality enriched environments, physical exercise and brain injury rehabilitation. IN Sharkey, P.M. (ed)*1st European Conference on Disability, Virtual Reality and Associated Technologies*, Maidenhead, UK, pp 247-252.

Johnson, D.A., Grealy, M.A., Rushton, S.K. & Soryal, I. (1997). Exercise and Virtual Reality : A new approach to rehabilitating patients with brain injury. *Journal of Neurotrauma.* 14, 253.

Johnson, D.A., Rose, D.F., Rushton, S., Pentland, B. & Attree, E.A. (1998). Virtual reality: A new prosthesis for brain injury rehabilitation. *Scottish Medical Journal,* 43(3), 81-83.

Johnson, D.A., Rushton, S.K. & Shaw, J. (1996). Virtual Reality enriched environments, physical exercise and brain injury rehabilitation. 1st World Congress in Neurological Rehabilitation, Newcastle, UK. (to be appear in *European Journal of Neurology,* 3)

Johnston, I.R., White, G.R. & Cumming, R.W. (1973). The role of optical expansion patterns in locomotor control. *American Journal of Psychology,* 86, 311-324.

Judge S.J. &Bradford, C.M. (1988). *Perception* 17, 783

Land, M.F. & Lee, D.N. (1994). Where we look when we steer. *Nature,* 369, 742-744.

Llewellyn KR. (1971). Visual guidance of locomotion. *Journal of Experimental Psychology* , 91, 245-261.

Longuet-Higgins HC,. Prazdny K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B* , 208: 385-397.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* W.H. Freeman and Co. San Francisco.

Mattingley, J.B., Bradshaw, J.L., & Bradshaw, J.A. (1994) Horizontal visual motion modulates focal, attention in left unilateral spatial neglect. *Journal of Neurology, Neurosurgery and Psychiatry,* 57, pp. 1228-1235.

McKee, S.P., Watamaniuk, S.N.J., Harris, J.M., Smallman, H.S. and Taylor, D.G. (1997). Is stereopsis effective in breaking camouflage for moving targets? *Vision Research,* 37, 2047-2056.

Milner A.D., Harvey M., Roberts R.C., Forster S.V. (1993). Line Bisection Errors In Visual Neglect - Misguided Action Or Size Distortion. *Neuropsychologia*,.31, 39-49.

McLeod, P., Driver, J. & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature,* 332, 154-155.

McLeod, P.D., McLaughlin, C. & Nimmo-Smith, I. (1985). Information encapsulation and automaticity: Evidence from the visual control of finely tuned actions. IN Posner, M. & Marin, O.S.M. (Eds). *Attention and Performance, Vol 11.* Erlbaum, Hillsdale, N.J.

Mon-Williams, M, Wann, J.P. & Rushton, S.K. (1993). Binocular Vision in a Virtual World: Visual deficits following the wearing of a head-mounted display. *Ophthalmic and Physiological Optics.* 13 , 387-391

Mollon, J. (1997). "....on the basis of velocity cues alone": some perceptual themes. *Quarterly Journal of Experimental Psychology,* 50A, 859-878.

Mon-Williams, M, Wann, J.P. & Rushton S. (in press) Investigating the reciporical cross-links between accommodation and vergence: Implications for virtual reality displays. *Ophthalmic and Physiological Optics.*

Nakayama, K. & Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature,* 320, 264-265.

Neeper SA, Gomez-Pinilla F, Choi J, Cotman C. (1995). Exercise and brain neutrophils. *Nature*, 373, 109

Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J., Goldstein, L., Heather, J., Frakowiak, R. & Frith, C. (1995). The physiology of coloured heared: A Positron Emission Tomography activation study of coloured-word synaesthesia. *Brain*, 118, 661-676.

Pizzamiglio, L., Frasca, R., Guariglia, C., Incoccia, C. & Antonucci, G. (1990). Effect of optokinetic stimulation in patients with visual neglect. *Cortex*, 26, pp. 535-540.

Pöppel, E. (1989) Taxonomy of the Subjective: An Evolutionary Perspective. IN Brown, J.W. (ed) *Neuropsychology of Visual Perception*. Lawrence Erlbaum Associates, London.

Pöppel, E. (1994). Temporal Mechanisms in Perception. *International Review of Neurobiology*, 37, 185-202.

Regan, D.M. & Beverley, K.I. (1982). How do we avoid confounding the direction we are looking with the direction we are moving? *Science*, 215, 194-196.

Regan D. & Beverley, K.I. (1979). Binocular and monocular stimuli for motion in Depth: Changing-disparity and changing-size feed the same motion in depth stage.*Vision Research*, 19, 1331-1342.

Renner MJ & Rosenzweig MR. (1987). *Enriched and impoverished environments : Effects on brain and behaviour.* Springer-Verlag, New York.

Riddell, P.M., Bullinaria, J.A. & Rushton, S.K. (in press). Modelling Adaptation in the Human Oculomotor System. *Proc. of Psychology and Neural Network Modelling Conference.*

Riddell, P.M. & Rushton, S.K. Neural network modelling of the development of accommodation and vergence. EPSRC Grant no. GR/L82274

Riddoch, G. (1917). Dissociation of visual perception due to occipital injuries, with especial reference to appreciation of movement. *Brain*, 40, 15-57.

Robertson, I.H., Tegner, R., Tham, K., Lo, A. & Nimmosmith, I. (1995). Sustained Attention Training for Unilateral Neglect - Theoretical and Rehabilitation Implications. *Journal of Clinical and Experimental Neuropsychology*, **17**, 416-430.

Rose, F.D. & Johnson, D.A. (1994). Virtual Reality in brain damage rehabilitation. *Medical Science Research*, 22, 82.

Royden CS, Banks MS, Crowell JA. (1992). The perception of heading during eye-movement. *Nature*, 360, 583-585.

Rushton, S.K. VREPAR: Stroke Modules report to EU.

Rushton, S.K., Johnson, D.A & Wann, J.P. (1996). Ecological Correlates of line bisection errors in unilateral visual neglect and their potential application. 1st World Congress in Neurological Rehabilitation, Newcastle, UK. (to be appear in *European Journal of Neurology*, 3)

Rushton, S.K., Coles, K.L. & Wann, J.P. (1996). Virtual reality technology in the assessment and rehabilitation of unilateral visual neglect. IN Sharkey, P.M. (ed)*1st European Conference on Disability, Virtual Reality and Associated Technologies*, Maidenhead, UK, pp 227-232.

Rushton S.K. & Corley M.M.B. (1998). The consistency of letter-colour pairings in three cases of synaesthesia. *Perception*, 27, Supplement, 13a.

Rushton S.K. & Corley M.M.B. (in prep). Mapping consistency in three cases of coloured-hearing.

Rushton S., Mon-Williams, M. and Wann, J. (1994). Binocular vision in a bi-ocular world: new generation head-mounted displays avoid causing visual deficit. *Displays* **15,** 255-260

Rushton, S.K. & Riddell, P.M. (in press) Developing Visual Systems and Exposure to Virtual Reality and Stereo Displays. *Applied Ergonomics*.

Segal, G.M.A. (1997). Synaesthesia: Implications for Modularity of the Mind IN Baron-Cohen, S. & Harrison, J.E (eds) (1997). *Synaesthesia: classic and contemporary readings.* Blackwell Publishers, Oxford, UK.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge University Press. Cambridge.

von Steinbüchel N.V & Pöppel E. (1993). Domains of rehabilitation: a theoretical perspective. *Behavioural Brain Research,* 56, 1-10.

Stone, S.P., Patel, P. & Greenwood, R.J. (1993). Selection of Acute Stroke Patients for Treatment of Visual Neglect. *Journal of Neurology, Neurosurgery and Psychiatry,* **56**, 463-466.

Walsh, V. (1996). Perception - The Seeing Ear, *Current Biology,* 6, 389-391.

Wann, J.P. (1996). Anticipating arrival: is the tau-margin a specious theory? *Journal of Experimental Psychology Human Perception and Performance.* 22, 1031-1048

Wann, J.P. Rushton, S.K. & Mon-Williams M. (1995) Natural problems for stereoscopic depth perception in Virtual Environments. *Vision Research.*, 35, 2731-2736.

Wann, J.P. & Rushton S.K. & Lee D.N. (1995) Can you control where you are heading when you are looking at where you want to go? In B. Bardy, R. Bootsma & Y. Guiard. (Eds) *Proc. of the 8th International Conference on Event Perception and Action.* 171-174

Warren, R. (1976). The perception of egomotion. *Journal of Experimental Psychology: Human Perception and Performance.* 2, 448-456.

Warren, W.H. (1995). Self-Motion: Visual Perception and Visual Control. IN Epstein, W. & Rogers, S. (eds). *Perception of Space and Motion ,* Academic Press, San Diego, 263-312.

Warren WH, Hannon DJ. (1988). Direction of self-motion is perceived from optical-flow. *Nature,* 336, 162-163.

Warren, W.H., Blackwell, A.W., Kurtz, K.J., Hatsopoulos, N.G. & Kalish, M.L. (1991). On the sufficiency of the velocity flow field. *Biological Cybernetics.* 65, 311.

Zeki, S. (1993). *A Vision of the Brain.* Blackwell Scientific Publications. London.

Chapter 2:

The *dipole* model of perception of time-to-contact

2.1 Foreword

2.1.1 Over-specification of time-to-contact

When a ball is approaching an observer, its time to contact with the face is specified by retinal size and rate of change of retinal size (Lee, 1976) and also by its absolute optic disparity (equivalent to required vergence angle) and rate of change of absolute optic disparity. Thus the visual system could potentially use one source, or a combination of both sources, of information. Alternatively, an observer could switch between sources on the basis of task demands or the quality of the information that each source provides.

Regan & Beverley (1979) pointed out that the relative usefulness (they termed it "effectiveness") of disparity and size (looming) is a function of the ratio of the size of the approaching object to the distance between the eyes (this is independent of distance to the object). For example, a fly is very small and so it is hard to pick up useful size and rate of change of size information due to the lack of sensitivity of the visual system. However, disparity is unaffected by the size of the fly. Sometimes an information source may be lost, for example if an eye is closed or occluded during the flight of an approaching object then disparity is no longer a useful source of information.

In the first case (with the fly) it would obviously make sense to base an estimate of TTC primarily on disparity. In the second case (with only one eye), to avoid errors in estimation of TTC it is necessary to switch to an estimate based exclusively on size. Thus, it would appear necessary to monitor both sources of information and switch emphasis between them.

The psychophysical data shows that human estimates of TTC are influenced by both disparity and size (Heuer, 1993, Gray & Regan, 1997). However a model for the combination of size and disparity in the estimation of TTC has not been proposed.

The model reported in the following paper was inspired by a dataset reported in Wann & Rushton (1995). At the time, we were able to describe the results - observers appear to use both disparity and size and place most emphasis on the cue that indicates earliest arrival (what we called most 'immediate'), but, we were unable to propose a mechanism that could account for the results. After a great deal of model building, I stumbled on the *dipole* model when I realised that a short-hand of using time or spatiotopic distance (metres from the observer) was obscuring a simple geometric relationship between size, disparity and distance: as an object gets closer, size and disparity do not increase linearly but almost exponentially (they are approximately inversely proportional to distance). The relationship is shown graphically in figure F1 below.

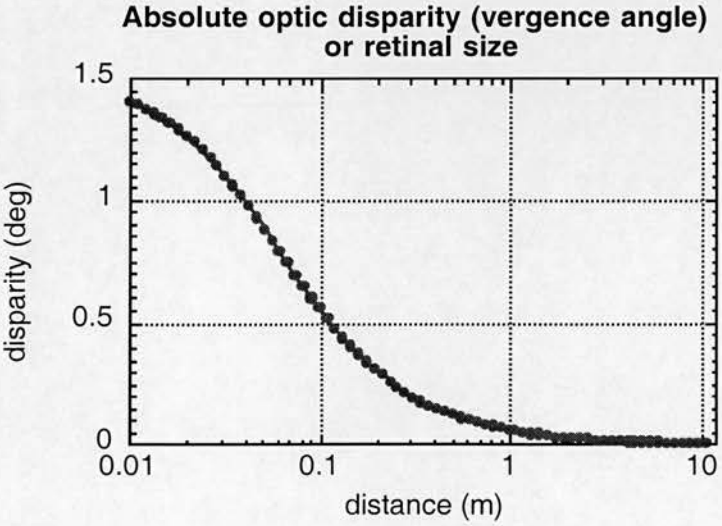


Figure F1 Absolute optic disparity (required vergence angle) calculated for an IPD of 62mm as a function of distance from the observer. This also corresponds to retinal size for a ball of 62mm diameter.

A trivially simple model that kept size and disparity in retinal units (distance on the retina or angular subtense) and combined them early (almost a classic tau-style model but taking disparity+size as input instead of just size) could account for the results. It also allowed some strong predictions to be made and tested. The paper that follows tells this story and presents data to support it.

2.2 Paper

from an early draft of Rushton, S.K. & Wann, J.P. (1999). Weighted Combination of Size and Disparity: A computational model of the timing of ball-catch. *Nature Neuroscience,* 2, 186-190.

2.2.1 Abstract

How do we time hand closure to cleanly catch a ball? Looming and changing binocular disparity provide the two primary sources of information about an object's motion in depth, but the relative effectiveness of either cue is dependent upon ball size. Following results from a virtual reality ball catching task, we derive a simple model that uses both cues and (i) is sensitive to the relative effectiveness of size and disparity, (ii) implicitly switches its response to the cue that specifies the earliest arrival and away from a cue that is lost or below threshold. We demonstrate the model's strength by predicting the response of participants to some very unusual ball trajectories.

2.2.2 Introduction

The time-to-contact (TTC) of an approaching object is specified by its current distance and rate of change of distance (velocity). The simplicity of this mathematical solution unfortunately does not reflect the scale of the problem involved in a neural or robotic implementation of TTC estimation. Survival in the forest, on the sports field or in the robotics lab, requires TTC be estimated very quickly and without undue call on cognitive resources. The use of distance and velocity requires construction of a 3D representation of the environment. This is non-trivial and demanding. To operate as fast as possible, therefore, a TTC mechanism may rely instead upon a simple sparse, or even crude, abstraction of the visual environment. It is also possible that in some settings specific sources of information may be lost, or fall below the threshold for detection (Regan & Beverley, 1979). The human performer seems to cope well with a range of environmental conditions arguing for a robust mechanism for TTC estimation that degrades gracefully in the case of cue-conflict or cue-loss.

2.2.3 Specification of TTC

Optical size and binocular disparity are both functions of the distance of the ball from the observer. Therefore, both size and disparity could provide estimates of TTC (Lee, 1976; Heuer, 1993; Regan & Hamstra, 1993; Gray & Regan, 1997). A significant number of studies of interceptive timing have ignored the role of disparity and have addressed the role of changing retinal size. Evidence for the exclusive use of retinal looming in natural TTC judgement, however, is weak (Wann, 1996). The relative amplitude of optic expansion ($d\theta/dt$) compared to changing disparity ($d\phi/dt$: Appendix Figure A1) is a function of object size relative to inter-pupillary distance and is independent of viewing distance (Regan & Beverley, 1979). Hence, depending on the size of the object (e.g. football vs. table-tennis ball), and the individual's sensitivity to looming or disparity, either input may be more salient at different stages of different tasks. Regan & Beverley (1979) provided a worked example where they suggested that the changing disparity of an approaching cricket ball should be much more salient than changing size for their observer TW. Heuer (1993) and Gray & Regan (1997), amongst others, provide empirical data that demonstrates that both changing disparity and changing retinal size support perception of TTC. Hence a model of a robust TTC system should take account of both changing size and changing disparity inputs (Regan & Beverley, 1979; Heuer, 1993; Regan & Hamstra, 1993) and ideally be sensitive to their relative rates of change ($[d\theta/dt]/[d\phi/dt]$).

Dissociation of TTC_{Disp} and TTC_{Size}

Under normal circumstances the TTC of an approaching object with the face and TTC with a hand placed on the ball's trajectory differ [TTC(Hand) = TTC(face) - Distance(hand from face)/speed(Ball) to a first approximation]. Further, the speed of an approaching object may not be constant, and the height of the ball will change during the trajectory. In the experiments described herein we use the following simplifications: There is a constant speed approach so that TTC and 'tau' (the TTC calculated using the current instantaneous speed) are the same; the ball does not change in height during approach; the hand is placed next to the face so that TTC(Hand) = TTC(Face); the ball always approaches down the mid-line, i.e. flies from a point directly ahead to a point between the eyes. These simplifications allow us to easily dissociate TTC_{Disp} and TTC_{Size}:

TTC is specified independently by both changing disparity (TTC_{Disp}) and by changing optic size (TTC_{Size}). The ratio of the instantaneous angular size to the rate of change of angular size ($\theta / (d\,\theta/dt)$) is proportional to the TTC with the face (when the object travels at a fixed speed), the ratio of the instantaneous disparity to the rate of change of disparity ($\phi / (d\phi/dt)$) is similarly proportional to the TTC. TTC_{Disp} and TTC_{Size} are dissociated as follows. Consider the approach of a solid ball straight towards the face: If the physical size of the ball is changed <u>during</u> flight then this manipulation does not change TTC_{Disp} ($\phi / (d\phi/dt)$) (We assume that mean disparity is used. If the front or back surface is used then this will introduce a very minor bias). However changing the size of the ball will change $TTC_{Size}(\theta / (d\theta/dt))$. For example, one could change the physical size so that the angular size does not increase as the ball approaches. in this case, $d\theta/dt$ would be zero.

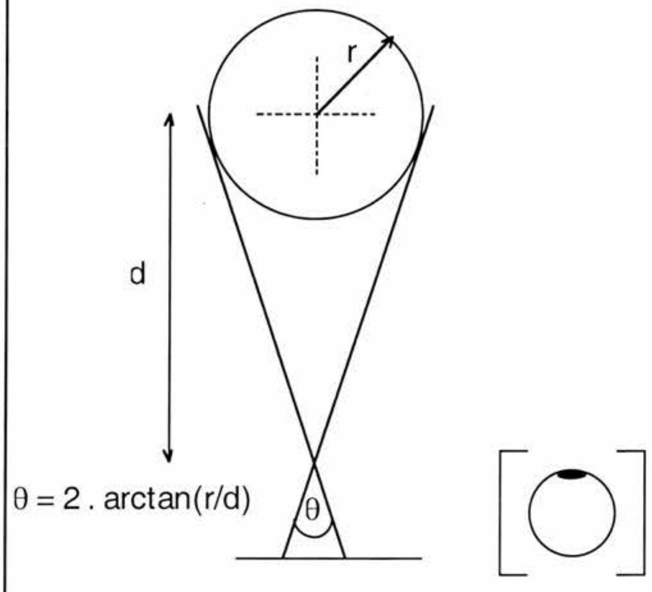


Fig X: The optic (angular) size of a ball before an eye. The above assumes that θ is small. Note that modifying r will change θ and hence (θ / (d θ/dt)).

It has been demonstrated that TTC judgements can be biased by binocular information. Some studies have used naturalistic tasks, with ball disparity manipulated by relatively crude devices (Judge & Bradford, 1988), whereas psychophysical settings have sometimes resulted in poor response accuracy (Heuer, 1993). Gray & Regan (1997), however, demonstrated that TTC judgements can be accurate to within 3-10% (51 - 270msec), for naive observers in a psychophysical task. Traditional psychophysics allows precise control over stimulus presentation, but raises the issue that it may not be valid to extrapolate the results to active tasks in naturalistic settings (Milner & Goodale, 1996; Tresilian, 1994). In order to retain the ability to precisely dissociate optical size and disparity, whilst keeping the task as a naturalistic action, we gave observers the task of catching tennis balls but within a computer generated virtual environment (Figure 1).

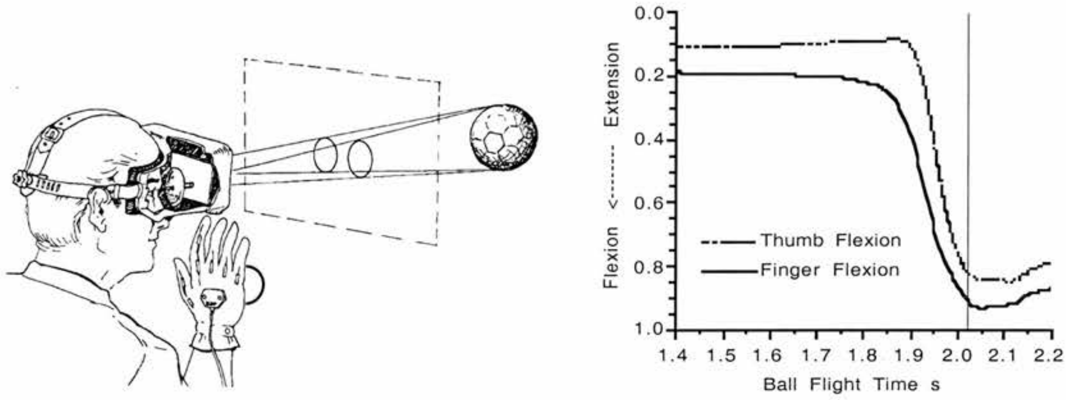


Figure 1: Left: Observers wore a head mounted display which placed a liquid crystal screen and magnifying optics in front of each eye to project stereoscopic images of a ball travelling towards them along a 5m stereoscopic corridor. **Right**: Finger and thumb flexion during the flight of the virtual ball. In the trial illustrated, the time of arrival of the ball at the eye would have been after 1.98s (start distance = 4.10m, velocity = 2.07m/s). It may be observed that the onset and completion of the grasp response appropriately bracket the time of ball arrival.

2.2.4 Experiment 1

Displays. Observers wore a head mounted display (HMD) which placed a liquid crystal screen and magnifying optics in front of each eye to project stereoscopic images onto an image plane. Horizontal display resolution was 220 colour pixels over 60 deg of visual angle per eye. Display update was >24 frames/s.

Procedure. Participants began each trial with their hand open by their face with a real tennis ball attached to their palm. A virtual stereoscopic textured tennis ball loomed towards them and they grasped onto the real tennis ball attached to their hand when they thought that the virtual ball would hit them. Finger and thumb flexion was recorded at 200Hz during the flight of the virtual ball. In both experiments the mean starting depth D_S was 4m, speed V was 2m/s and initial ball diameter R_S was 7cm. D_S, V and R_S were randomly varied by ± 10% across trials.

The HMD was fully immersive and ball flew down a stereoscopically presented 5m corridor towards the face of the observer.

Participants (n = 6) received 20 practice trials. A set of 20 control trials (Condition A) were then interleaved with 20 trials where looming was computationally scaled to provide a TTC estimate 100ms earlier (Condition B) or 100ms later (Condition C) than that specified by binocular disparity. A virtual image of the moving hand was not available to participants, to avoid inter-trial learning.

2.2.5 Results, Experiment 1

The initial finding was that observers were able to accurately catch virtual balls. Their grasp response in the control condition was initiated 146msec before the ball arrived (se = 15msec) and completed 22msec after the ball arrived (se = 14ms) (see figure 2 below). This accuracy is equivalent to that required in natural catching (Alderson, Sully & Sully, 1974).

When looming indicated a different TTC from disparity observers grasped significantly earlier or later, respectively. The change in grasp time, however, was not equal across the two conditions (Figure 2).

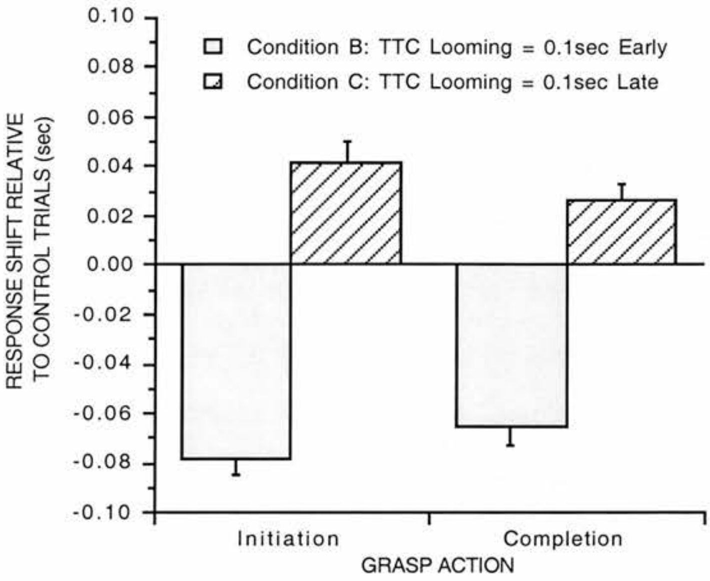


Figure 2: Changes in the time of grasp initiation and completion when participants (n = 6) were catching virtual balls for which looming was scaled to provide a TTC estimate 100msec earlier (Condition B) or 100msec later (Condition C) than the binocular TTC information. The unsigned change in grasp time was significantly greater in Condition B (gain for looming = 0.78) than Condition C (gain for looming = 0.41) for both initiation and completion (t(5) = 3.51 , p <.01; t(5) = 3.34, p<.01, one-tail, respectively).

Hence, the results demonstrate that both optical size and disparity are used in interceptive actions, but the respective influence or 'weighting' of size and disparity is not fixed. Looming was the dominant cue when it specified a TTC 100msec earlier than disparity (Condition B), whereas disparity had the most influence when looming specified a later TTC (Condition C). The earliest, (what we will term the most 'immediate'), estimate of arrival had the greatest influence on grasp timing.

2.2.6 Weighted Models

To account for the findings we need to derive a model that uses both size and disparity and automatically biases response towards the most immediate cue. We identify three potential models that combine optic size, disparity, rate of change of size and rate of change of disparity to estimate TTC.

The first model is a fixed weighting modular summation (Fig 3A). In this model two separate estimates of TTC are derived from (i) size and its rate of change and (ii) disparity and its rate of change. The two estimates are then combined according to some weighting factor β, where $\beta=1$ would reduce the model to one that relied exclusively on size (Lee, 1976), and $\beta=0$ would use only disparity. Although this model includes both size and disparity, its fixed weighting precludes it from biasing its TTC estimate towards the most immediate cue and so it cannot account for the flexible response demonstrated in our data. It would also have difficulty coping with sudden cue-loss.

The second, a variant of the first, introduces changes in the β weighting "on the fly" on the basis of relatively immediacy of the two modular estimates (TTC_{Size} & TTC_{Disp}):

$$TTC = \beta_1 * TTC_{Size} + \beta_2 * TTC_{Disp}, \qquad [1]$$

This weighting system would be robust to cue loss and account for the results in Fig. 2 if β_1 ,β_2 were dynamically revalued using cross ratios of the modular inputs (Fig 3B):

$$\beta_1 = TTC_{Disp} / (TTC_{Size} + TTC_{Disp}) \qquad \beta_2 = TTC_{Size} / (TTC_{Size} + TTC_{Disp}) \quad [2]$$

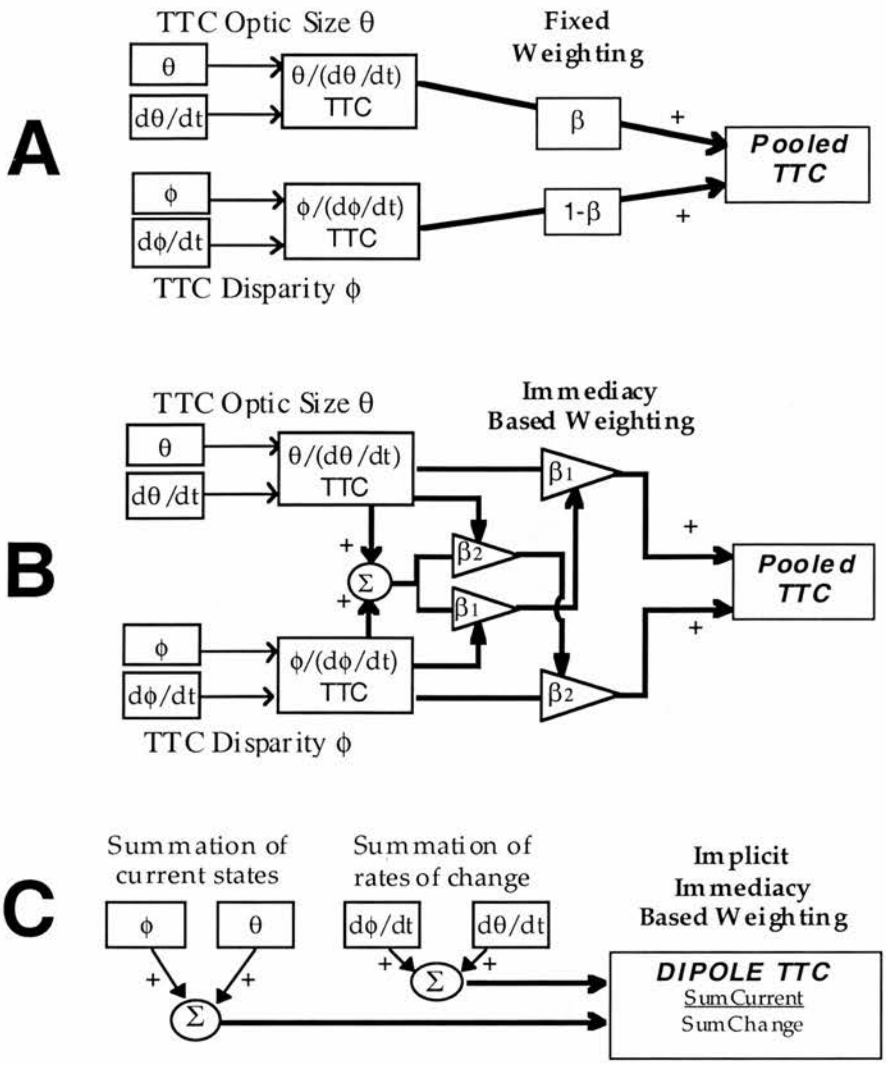


Figure 3: Models that combine the information from binocular and monocular sources of time to contact (TTC). **A. Fixed weight modular combination**: TTC is estimated from each input and then averaged on the basis of some weighting (0< β <1). This approach, however, is not robust if one input is perturbed or lost, because TTC = β θ/(dθ/dt) + (1-β) ϕ/(dϕ/dt) will produce large errors of estimation. **B. Flexible weight modular combination:** A variant of model A where each modular estimate is first summated (Σ) and then a cross ratio from the opposing input used to calculate the β weightings (see main text). **C. Dipole Combination:** binocular and monocular motion is combined prior to estimating TTC. In the case of perturbation or drop out of one of the inputs, this approach will produce a TTC estimate that is always biased towards the cue specifying the earliest arrival, thereby exhibiting the advantages of model B without the explicit iterative calculation of β weights in flight.

The third model (Fig 3c) does not calculate TTC separately for each input, but combines instantaneous size and disparity and uses the summed parameter to calculate TTC. Two points (from a texture or perimeter) viewed by a single eye specify a *dipole* (Julesz, 1981), and a single point viewed by a pair of eyes can also be considered a *dipole*, due to geometric equivalence (Appendix Figure A1). Our computational model treats an object as a collection of *dipoles*. The *dipole* extents are summed and the rate-of-change of *dipole* lengths are either combined from local detectors (Eq. 3) , or derived from the change in *dipole* length (Eq. 4).

$$TTC_{Dipole} = (\theta + \phi) / (d\theta/dt + d\phi/dt) \qquad [3]$$

'Dipole'

Julesz introduced the term *dipole* to the theory of texture perception when describing the statistical distribution of texture elements (a *dipole* is the second order statistic). A *dipole* is an extent defined by two end-points. For example, a needle is a *dipole*. The term *dipole* was chosen to describe the model of TTC primarily for historical reasons. I had been contemplating an implementation of the model in terms of a summed speed field. If a textured object is approaching, each texture element defines a velocity vector. The model I was considering took no notice of the directional component of the velocity vectors just their magnitude. Trying to describe the model in terms of the 'speed component of the velocity vector' is very clumsy and the term dipole served to replace such an awkward formulation. The issue of implementation was forgotten but the term stuck. The choice seems useful as the issue of implementation has still to be resolved and describing the input to the model in terms of *dipoles* serves a very important purpose - to indicate that the model does not distinguish between binocularly derived and monocularly defined input.

The retinotopic quantities $(\theta + \phi)$ rise non-linearly as the object approaches. As a result, the cue that specifies the earliest arrival has a disproportionate influence. For a non-deformable object of diameter R and inter-pupilary distance I, equation 3 is equivalent to:

$$TTC_{Dipole} = (TTC_{Size} + TTC_{Disp} [I/R])/ (1 + [I/R]) \qquad [4]$$

Hence when I = R the *dipole* model bases its estimate on an equal weighting of both inputs, but for a football the implicit bias would be towards an estimate based on TTC_{Size}, and for a table-tennis ball towards TTC_{Disp}.

If one input is lost (e.g. $\phi = 0$ or $\theta = 0$) then the estimate relies upon the remaining single input (Eqn 3). In the case of a rate of change being lost or below threshold, the model predicts a relatively small response error (Appendix: A12). If $\phi>0$ and $\theta> 0$, but changing size specifies a different arrival time to changing disparity, then the *dipole* model biases its estimate towards the most immediate cue equivalent to the explicit weighting of equation 2 (Fig. 4 upper). In the case of the cue-manipulation employed in Experiment 1 the *dipole* model predicts associated temporal shifts of -70msec when size specifies an earlier arrival and +34 msec when size specifies a later arrival. This is in good agreement with the mean results presented in Figure 2. The *dipole* model is compatible with psychophysical results suggesting that changing size and changing disparity feed a common motion-in-depth system (Regan & Beverley, 1979). It is also consistent with the results of other researchers who have presented cue conflicts. Heuer (1993) reported TTC judgements for conditions where TTC_{Size} and TTC_{Disp} were either 4sec or 8sec, but placed in conflict. The dipole prediction from the initial conditions would be a TTC estimate of 5.3sec, which is similar to the performance observed.

2.2.5 TTC from relative disparity

An accurate geometric specification of TTC from binocular information requires the detection of the binocular angular subtense (absolute disparity, Howard & Rogers, 1995). It has been demonstrated, however, that a change in absolute disparity (ϕ) produces no sensation of motion in depth (Regan, Erkelens, Collewijn, 1986) and motion is only seen when there is a reference for relative disparity (α, Appendix: A2, A3). Provided there is a stable structured optic array, which there is in most natural environments, then $d\alpha/dt$ is equivalent to $d\phi/dt$ and this is not an issue, but the use of α in place of ϕ predicts systematic errors (Appendix: A13, A14). It has been proposed that spatiotopic distance (D) could be used in combination with $d\alpha/dt$ to estimate TTC (Regan & Hamstra, 1993), but estimation of absolute distance (D) for a rapidly moving object is a considerable problem (Regan, 1992). It has also been demonstrated that observers can estimate TTC from disparity when all cues to D are removed except for relative disparity and fixation vergence (Gray & Regan, 1997). The latter finding supports an argument that TTC judgements were based on $\alpha/(d\alpha/dt)$ or that $\phi/(d\alpha/dt)$ was recovered from fixation vergence (Appendix: A4, A5). In our experiments, relative disparities were available from reference objects, and we assume that α and $d\alpha/dt$ were used in place of absolute disparity. This predicts a timing error with the dipole model of approximately 18msec early for Experiment 1 (Appendix: A15, see also Figure 4 upper). Perhaps most important, is that our model copes with the dropout of relative disparity, such as for a high ball where a reference point may be lost. In this case equation 3 reduces to a single cue estimate based on optic size and then implicitly switches back to weighting disparity if a reference for its detection becomes available as the ball drops.

2.2.6 Pitching a tricky ball

The dipole model was built to account for the findings outlined in Fig. 2. Although we have argued the model is compatible with previous psychophysical data, we wished to subject the dipole model to a stronger empirical test. We have outlined that the model can cope with complete loss of one cue or a static state for one input (Appendix: A10, A11). The "stalling" of one TTC estimate, however, would predict a specific response delay (Fig. 4 upper).

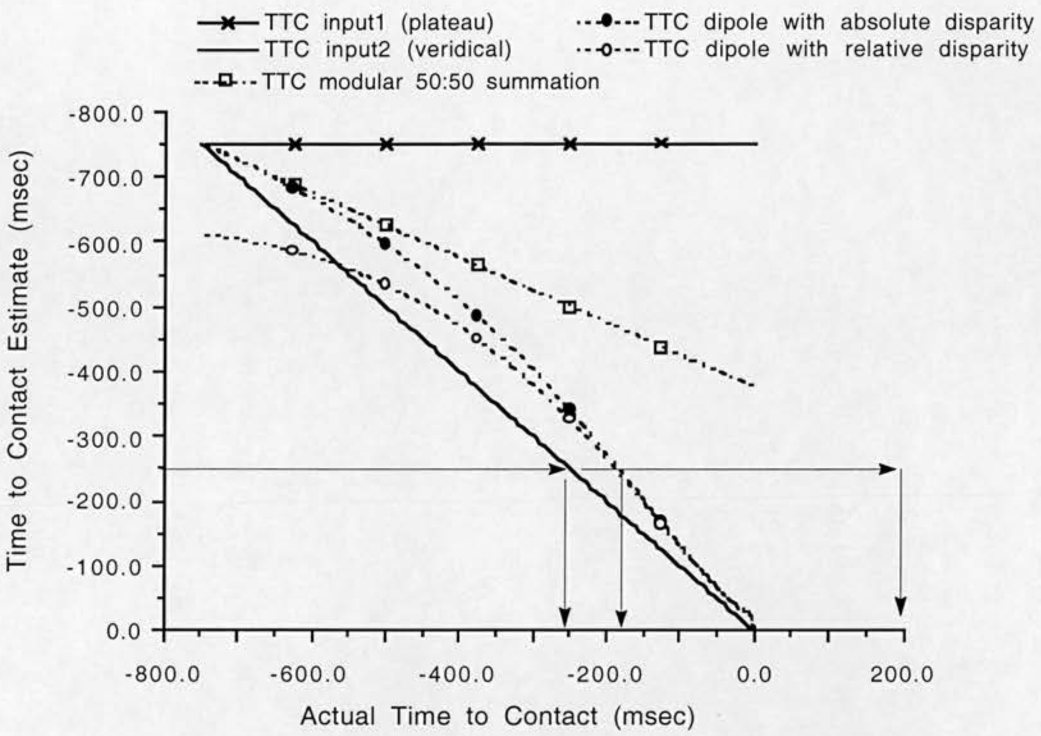


Figure 4. TTC estimates from a dipole and modular model for the TTC plateau conditions. Solid lines indicate the potential inputs from changing size and changing disparity, where input_1 indicates a constant TTC of 750msec and input_2 is veridical. The modular estimate (open squares) is based on a 50:50 weighting of both inputs, whereas the dipole (filled circles: absolute disparity) displays an increasing bias towards the most immediate input. Open circles illustrate the dipole TTC estimate when using relative disparity for these experimental conditions with a reference point at 4m. Arrows illustrate that if the desired response time was 250msec TTC, then the dipole would predict a response delay of 70ms, where the modular model would predict a response well after the ball had arrived. The error in using relative disparity is minor.

We computationally scaled either the optical size or disparity so that they produced a TTC plateau. Hence, in one condition changing size presented a normal decreasing TTC estimate, but disparity indicated a constant TTC, whereas the reciprocal condition produced decreasing TTC for disparity and constant TTC for size (Figure 4 upper). In both cases the perturbations do not "freeze" the ball in disparity or size. Optic size and disparity continue to change, but their temporal derivatives are decreased at a rate that keeps TTC constant.

2.2.7 Methods, Experiment 2

Methods were the same as for the previous experiment (2.2.4), except for the following details: The HMD was see-through and allowed the image of the ball to be overlaid onto the natural world, that contained edges at a range of depths. Horizontal resolution was 263 colour pixels over 30 deg of visual angle per eye. The vertical field of view was 45 and 23 deg respectively. Display update was >24 frames/s. The screen edges of the display were also fusable at approximately 4m, for someone with an IPD of 65mm. The system was calibrated to individual participants using a static image of the ball at 3 different distances. Participants (n = 5) held their hand next to their cheek so it was occluded from their vision and received 20 practise trials as in experiment 1. A set of 20 control trials were then interleaved with plateau trials, where the ball moved identically to the control trials until it was 1.5m from the participant and then either optic size was scaled to keep $\mathrm{TTC}_{\mathrm{Size}}$ constant at 750msec or disparity was scaled to keep $\mathrm{TTC}_{\mathrm{Disp}}$ constant at 750msec while the other input continued to decrease in line with the control trials.

2.2.8 Results, Experiment 2

So how might an observer cope with the strange situation of a ball where one cue indicated that the ball will arrive soon and the other cue indicates that it is moving forward, but will never arrive?

Figure 5 (hatched bars) suggests that when disparity TTC plateaus, observers relied almost exclusively on looming information. But in contrast, if the next ball had the TTC plateau occurring on looming, then observers switched their estimates to disparity. Stimulus dependent switching of this type is a basic feature of the dipole model and the predictions of the dipole model are remarkably similar to the mean tendency across participants. When both inputs were scaled to a similar plateau, the observer simply saw a ball slowing down and there was a long delay in their response.

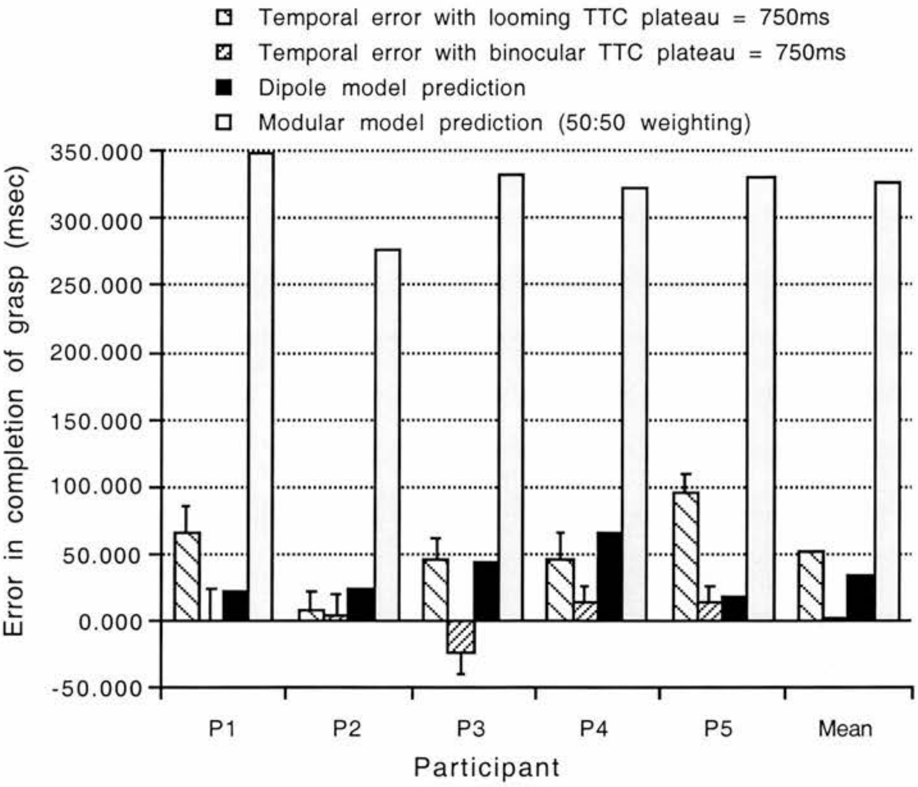


Figure 5: Results of a replication of the paradigm presented in Figure 1, but with scaling of the virtual images to provide a TTC plateau for either optic size or disparity. In the plateau trials, optic size was scaled to keep TTC_{Size} constant at 750msec while TTC_{Disp} continued to decrease in line with the control; or TTC_{Disp} was kept constant at 750msec, while TTC_{Size} continued to decrease in line with the control trials. When there was a disparity plateau (hatched) there were only small response errors indicating that most weight was attached to the optic size estimate, whereas the converse is true when there was a looming plateau. Predictions from the dipole model and the fixed weighting model were produced by using the response time for the control trials for each participant to estimate when that TTC would be achieved for each model. The fixed weighting model predicts large errors not seen in this experiment.

2.2.9 Summary

Research into binocular vision has predominantly focused on the role of relative disparities (Regan & Beverley, 1979; Gray & Regan, 1997) and their role in relative depth, and motion in depth, judgements. Catching a ball in flight requires more than just disparity and its rate of change and seems to be based upon a weighted estimate of changing size and changing disparity. The dipole model allows for the influence of both optic size and disparity in line with our experimental findings and is compatible with previous findings. It implicitly compensates for changes in the relative effectiveness of looming or binocular motion as a function of object size (Regan & Beverley, 1979). It also displays two ecologically useful features: in the case of cue-conflict it biases the TTC estimate towards the cue that signals the earliest arrival; if a cue is lost then the model exhibits graceful degradation by switching to the remaining source of information. In this respect it is the first computational account of how the visual system may handle sensory conflicts that arise from inputs that are below motion threshold and therefore indicate an infinite TTC. Cue switching is an implicit feature of the basic architecture of the dipole model and as such the model provides a simple appealing account of adaptive behaviour. It is possible to design more cumbersome modular architectures. However, compatibility with Regan & Beverley's (1979) motion-in-depth stage, simplicity and efficiency favours the dipole model.

2.2.11 Mathematical Appendix.

This section lays out the mathematical details of the dipole model. I was responsible for the basic model. I developed the model and worked predictions via simple computational simulations. The formal derivations laid out below were predominantly the work of John Wann. I include the section here for completeness.

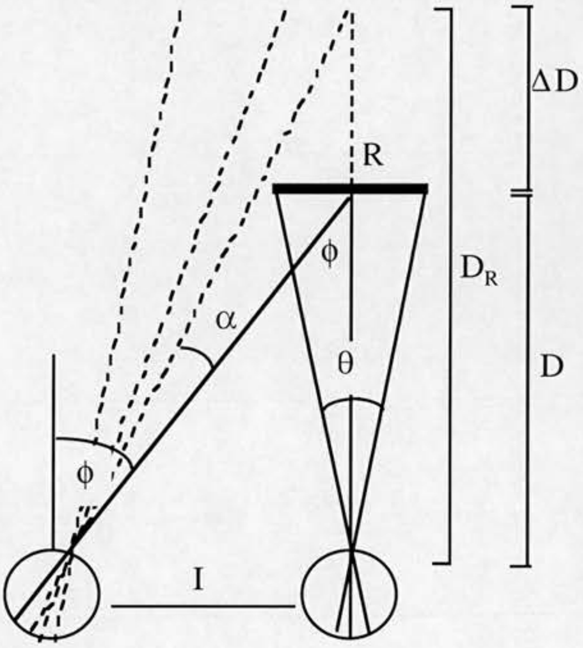


Figure 5

Taking the case in Fig. 5 of a ball of diameter R travelling at a constant velocity V towards an observer who has an interpupilary distance I. Broken lines represent the subtense of reference objects in the environment. For $D >> I$ and $D >> R$ optic size $q \approx R/D$ and absolute optic disparity $\phi \approx I/D$. Their derivatives can be approximated by:

$(d\theta/dt) \approx RV/D^2$

$$(d\phi/dt) \approx IV/D^2 \qquad (A1)$$

Relative disparity can be approximated by

$$\alpha \approx I\Delta D/D(D+\Delta D) \qquad (A2)$$

Which for an approaching ball does not reduce to the normal expression of $I\Delta D/D^2$

The derivative of A2 is identical to A1:

$$(d\alpha/dt) \approx IV/D^2 \qquad (A3)$$

Given these approximations there are a number of issues that arise in the specification of time to collision (TTC) from monocular and binocular inputs:

Using spatiotopic estimates in TTC judgements

It has been proposed Gray & Regan (1997) that performers could base their judgements of TTC upon:

$$TTC = I / [D(d\alpha/dt)] \qquad (A4)$$

There are very few cues to absolute distance of an object in flight, however, and the estimation of D is likely to be difficult (Regan, 1992). We note that for small optical angles, the spatiotopic distance D of the object could be estimated from retinotopic angles using:

$$D = I\,(d\theta/dt)/[\theta(d\alpha/dt)] \qquad (A5)$$

But substituting A5 into A4 reduces the estimate to a single source retinotopic estimate of $\theta/(d\theta/dt)$. Hence, although it is possible, in principle, to recover the spatiotopic parameters such as D from retinotopic information this does not support an argument that spatiotopic estimates are used. In reduced cue environments (Gray & Regan, 1997) an estimate of D would need to arise from extra-retinal information and there is no clear mechanism for supporting this approach.

Behaviour of the dipole model for scale changes and cue-loss conditions

The geometric specification of the dipole model bases its TTC estimate upon:

$$\mathrm{TTC}_{\mathrm{Dipole}} = (\theta + \phi) \,/\, [\,(d\theta/dt) + (d\phi/dt)] \qquad \text{(A6)}$$

This can be rewritten for the case where $(d\phi/dt) > 0$ and $(d\theta/dt) > 0$ to demonstrate that the model biases its estimate towards whichever expansion pattern is most salient:

$$\mathrm{TTCdipole} = \frac{\mathrm{TTCSize} + \mathrm{TTCDisp}\,\dfrac{(d\phi\,/\,dt)}{(d\theta\,/\,dt)}}{1 + \dfrac{(d\phi\,/\,dt)}{(d\theta\,/\,dt)}} = \frac{\mathrm{TTCSize}\,\dfrac{(d\theta\,/\,dt)}{(d\phi\,/\,dt)} + \mathrm{TTCDisp}}{\dfrac{(d\theta\,/\,dt)}{(d\phi\,/\,dt)} + 1} \qquad \text{(A7)}$$

Substituting from A1 gives an approximation for the ratio of object size R to interpupilary distance I

$$\mathrm{TTC}_{\mathrm{Dipole}} = (\mathrm{TTC}_{\mathrm{Size}} + \mathrm{TTC}_{\mathrm{Disp}}\,[I/R]\,)/\,(\,1 + [I/R]\,) \qquad \text{(A8)}$$

Which conveniently adjusts the weighting for the TTC estimate in line with the relative strength of changing size and changing disparity noted by Regan & Beverley (1979):

$$R/I \approx (d\theta/dt)/(d\alpha/dt) = (d\theta/dt)/(d\phi/dt) \qquad \text{(A9)}$$

If one input is lost (e.g. occlusion of one eye or no reference for relative disparity) then A6 and A7 reduce to a single input estimate:

$$\mathrm{TTC}_{\mathrm{Dipole}} = \theta/(d\theta/dt) \qquad \text{(A10)}$$

It may be the case, however, that one of the of the derivatives in A6 is below threshold (e.g. for a small object where $d\phi/dt$ is lost) then $TTC = \infty$ for that input and this will introduce an error term:

$$TTC_{Dipole} = \theta/(d\phi/dt) + \phi/(d\phi/dt) \qquad \text{(A11)}$$

For these cases A10 can be rewritten to demonstrate that the error would be equivalent to the ratio of two angular subtenses:

$$\text{TTCdipole} = \frac{\text{TTCSizeTTCDisp}\,(1 + \theta / \phi)}{\text{TTCSize} + \text{TTCDisp}\,(\theta / \phi)} \qquad \text{(A12)}$$

If the thresholds for motion detection are of a similar order for both $(d\theta/dt)$ and $(d\phi/dt)$ then the errors should only occur when $\theta \ll \phi$ (e.g. $R \ll I$) or $\phi \ll \theta$, hence the dipole model predicts a minor error. In some viewing conditions with flat-screen stimuli, the size or disparity of a target is deliberately held to be static and this information presents a TTC input of infinity (Regan & Beverley, 1979; Gray & Regan, 1997). For that case the error would be equivalent the ratio of static/moving angular subtense and the results of Heuer (1993) and the static size results of Gray and Regan (1997) are similar to these predictions.

Using relative disparity as an estimate of ϕ

A change in absolute disparity ϕ, does not result in a percept of motion in depth, unless there is a concomitant change in relative disparity α (Erkelens & Collewijn, 1985). The rate of change of ϕ and α have geometric equivalence (A1, A3) hence the issue is the use of α instead of ϕ in estimating actual TTC. Using A1 and A2:

$$\alpha/(d\alpha/dt) = TTC(\Delta D/(D+\Delta D) = TTC - TTC(D/D_R) \qquad \text{(A13)}$$

The error term of TTC(D /D_R) can be substituted into A8 to predict an error for the dipole of:

$$TTC_{Dipole\_Error} = -[\,TTC\,(I/R)\,(D/D_R)]/(1 + [I/R]) \qquad (A14)$$

For our case of a cricket or tennis ball where $R \approx I$ this reduces to

$$TTC_{Dipole\_Error} = -0.5\,TTC\,D/D_R = -0.5\,TTC^2\,V/D_R \qquad (A15)$$

Hence early in the approach A15 specifies an error of early arrival (e.g. the catcher begins to orient earlier than required), then as $D \ll D_R$ the error is reduced as a function of TTC^2. If the final pick-up of visual information for catching is 300msec before contact (Gray & Regan, 1997), then for a ball travelling at 5m/s and $D_R = 10$m the error would be -22msec. If more distant reference points are available the error is negligible, but it is also directly scaled with the ball speed. The most difficult case is a small fast moving object, such as a table tennis ball where $I = 3R$. In this case, binocular information should be a primary source of information (A9), but the errors in TTC estimation may also be higher (A14). This is not at odds with the observation that such skills require considerable practise and undoubtedly part of what is learnt are minor timing adjustments in the hitting/catching response.

In summary relative disparity, α, can provide a useful estimate of TTC, provided the final adjustments of the interceptive action are late in the ball trajectory. It is also the case that the weighting of changing size in the dipole model reduces the error in using relative disparity by a factor of $(I/R)/(1+I/R)$.

Moving away from the eye

In experiment 1 and 2 we set participants the task of catching a ball next to their face. This was because the derivations of TTC estimates in this and previous papers (Lee, 1976; Gray & Regan, 1997; Regan & Hamstra, 1994; Wann, 1996) are only true for collision with the plane of the eye. Hence, if participants exhibit errors in "natural" catching some distance in front of the face, it is difficult to ascertain if these errors are implicit to the model or are due to errors in extrapolating their eye-point estimate, or errors in hand positioning. The model we present can be extended to cope with collision away from the plane of the eye. It has been proposed that balls with lateral motion may be caught eccentric to the head by combining optic expansion with the optical gap between the ball and its future catching location (Bootsma & Oudejans, 1993; Peper, Bootsma, Mestre & Bakker, 1994). Using the dipole model in place of the optic expansion makes this solution more robust for small objects. For head on approaches where the ball is to be caught some distance D_{hand} in front of the head, the participant must recover:

$$TTC_{hand} = TTC_{eye} - TTC_{eye}(D_{hand} / D_{ball}) \qquad \text{(A16)}$$

This can be approximated by using the disparities of the hand and ball relative to a reference object, which predicts an early arrival error when the ball is distant, but converges on a veridical estimate as the ball approaches:

$$TTC_{hand} = TTC_{eye}\,(1 - \alpha_{ball}/\alpha_{hand}) \qquad \text{(A17)}$$

2.3 Afterword

2.3.1 Information sources and the *dipole* model

So how do we categorise the *dipole* model in terms of information sources? The way we have conceptualised it, the *dipole* model relies on a single source of information, *dipoles*. However, a traditional analysis of the information available in the binocular array would separate out binocularly defined and monocularly defined *dipoles*.

The *dipole* model is sensitive to more than one type of information by a classical definition (disparity and size). Further, it "switches between" information sources: when an object is small it relies on disparity, when one information source specifies an earlier arrival it switches towards that source. However, the switching between information sources simply follows from geometry, there are no switches within the model.

2.3.2 Strong versus weak fusion

Depth is specified by many different cues. Landy et al (1995) discuss "strong" vs. "weak" (Yuille & Bülthoff, 1995) fusional models of depth. "Strong" vs. "weak" can be thought of as a question about granularity of modularity.

In the case of depth, Landy et al explain that perception of depth may rely on either a combination of estimates from a series of systems each dedicated to a different depth cue (for instance, one module for stereo depth, one module for texture defined depth etc.). This would be *weak* fusion.

Alternatively, perception of depth could rely on a single depth system that does differentiate different depth cues and process them separately. This would be *strong* fusion.

By this criterion the *dipole* model would be classified as a strong fusion model. Disparity and size are not processed individually to produce independent estimates of TTC. Instead both are used in a single system.

One consequence of this is that it should preclude the possibility of changing the weighting or influence of each cue. This matter is addressed in more detail in the following chapter.

2.3.3 Implementation - flow components?

How might the *dipole* model be implemented? The model could be implemented using traditional binocular information, disparity and rate of change of disparity. Alternatively it could be implemented with flow components. A flow implementation could work in a monocular or Cyclopean frame without disparity (ie at no point would it be necessary to correlate left and right half images and solve the "correspondence problem".

If gaze is fixed, when an object moves in depth towards an observer, it also moves nasal to temporal across the retinas. If we think of this movement in terms of flow then it is 'translational' flow. The rate of translational flow can be substituted for rate of change of disparity in the *dipole* model. Thus the dipole model could simply combine translational and radial flow (expansion or looming).

It has long been recognised that the mechanism for motion-in-depth could be based upon use of relative left and right eye image velocities rather than changing disparity.

An influential paper by Cumming & Parker (1994) had been taken to settle the issue in favour of the use of disparity in perception of motion-in-depth. However, the topic has recently has recently been re-examined (Allison, Howard & Howard, 1998). At the moment the balance of evidence would remain in favour of changing disparity. However, given that there is support for different mechanisms for perception and action (Milner & Goodale, 1996) and Zihl & von Crammon's famous motion blind patient can catch a ball with the skill of a woman of her age (reported by Peter MacLeod and cited in Weiskrantz, 1997), presumably without experiencing a percept of motion-in-depth, it is sensible not to assume the same mechanism serves the percept of motion-in-depth and the system for perception of TTC.

An implementation of the *dipole* model in terms of flow components (radial flow + translational flow) that avoids use of disparity is aesthetically pleasing. However, this is not good enough reason to assume such an implementation is likely. However, I do have some anecdotal evidence to suggest that it is a possibility: an observer in one of the virtual ball-catching studies was unable to see depth in TNO stereo plates (standard tests of stereo acuity) but showed biases in response following from manipulation of disparity similar to those demonstrated by other observers.

A potential way to investigate this matter empirically would be to take advantage of individual differences: In the experiments reported in this chapter it can be seen that there is some difference between individuals. Some individuals appear to rely more on either looming or disparity. An individual's sensitivity to disparity, radial and translational flow can be measured. By looking at correlations between these measures and performance on the ball-catch task it may be possible to gain some insight into mechanisms.

2.4 References

Alderson G.J.K. Sully D.J. & Sully H.G. (1974). An operational analysis of a one-handed catching task using high-speed photography. *J. Motor Behav,* 6, 217-226.

Allison, R.S., Howard, I.P. & Howard, A. (1998). Motion in depth can be elicited by dichoptically uncorrelated textures. *Perception*, 27, Supplement, 46b.

Cumming, B. & Parker A.J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research,* 34, 483-496

Cutting, J.E. (1986). *Perception with an eye for motion.* MIT Press, London.

Gray R. & Regan, D. (1997). Accuracy of estimating time to collision using binocular and monocular information.*Vision Research,* 38, 499;

Heuer, H. (1993). Estimates of time to contact based on changing size and changing target vergence. *Perception* 22, 549.

Judge S.J. & Bradford, C.M. (1988). *Perception* 17, 783

Julesz, B . (1981). Textons, the elements of texture perception and their interactions. *Nature* 290, 91-97.

Laurent, M., Montagne, G. & Durey, A. (1996). Binocular invariants in interception tasks: a directed perception approach. *Perception*, 25, 1437-1450.

Lee, D.N. (1976). A theory of visual control of braking based on information about time-to-collision *Perception,* 5, 437-459.

Milner A.D. & Goodale, M.A. (1996). *The visual brain in action.* Oxford University Press, UK.

Regan, D. (1992). Visual judgements and misjudgements in cricket, and the art of flight. *Perception* 21, 91.

Regan D. & Beverley, K.I. (1979). Binocular and monocular stimuli for motion in Depth: Changing-disparity and changing-size feed the same motion in depth stage.*Vision Research,* 19, 1331-1342.

Regan, D., Erkelens C.J. & Collewijn H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology and Vision Science,* 27, 584-59.

Regan D. & Hamstra, S.J. (1993). Dissociation of discrimination thresholds for time to contact and for rate of angular expansion.*Vision Research,* 33, 447

Tresilian, J.R. (1995). Perceptual and cognitive processes in time-to-contact estimation: Analysis of prediction-motion and relative judgment tasks. *Perception & Psychophysics,* 57, 231-245.

Wann, J.P. (1996). Anticipating arrival: is the tau-margin a specious theory? *Journal of Experimental Psychology Human Perception and Performance.* 22, 1031-1048

Wann, J.P. & Rushton S.K. (1995). Grasping the impossible: Stereoscopically presented virtual balls In B. Bardy, R. Bootsma & Y. Guiard. (Eds) *Proc. of the 8th International Conference on Event Perception and Action.* 207-210

Weiskrantz, L. (1997). *Consciouness lost and found.* Oxford University Press. Oxford.

Yuille, A.L. & Bülthoff, H.H. (1995). Bayesian decision theory and psychophysics. IN Knill, D.C. & Richards, W. (eds) *Bayesian Perspectives on visual perception.* Cambridge University Press, Cambridge.

Chapter 3:

Can the influence of disparity be altered for interception in front of the face?

3.1 Forward

The purpose of this chapter is to follow up one of the points raised in the afterword of the previous chapter, Chapter 2. As noted, if perception of TTC relies on strong fusion then it should not be possible to change the "gain", or influence, of looming and disparity. The experiment that follows was designed to create circumstances conducive to an observer learning to ignore disparity. The experiment and results are reported in this short chapter. The data suggest that it is not possible to learn to ignore disparity. Hence strong fusion is a better model than weak fusion (Landy et al, 1995) for the interaction of disparity and looming in the perception of TTC before the face. It was not the purpose of this experiment to attempt to distinguish between various possible implementations that may underly perception of TTC and adaptation. However, for reference, I lay out candidate models in a lengthy appendix.

This chapter does not derive from a paper. The data was collected several years ago with Mark Mon-Williams and John Wann. The dataset needs to be extended before many of the interesting questions about implementation can be properly resolved.

To ensure the format of this chapter is consistent with the rest of the thesis I have written it in paper form.

3.2 Paper

3.2.1 Abstract

Intercepting an approaching object requires a judgement of time-to-contact (TTC). Both binocular disparity and looming influence TTC judgements (Gray & Regan, 1997) and when disparity is manipulated with telestereoscopes, then the ability to catch a ball is disturbed (Judge & Bradford, 1988). It has been shown in studies concerned with judgements of depth that when one cue is made unreliable, its influence is diminished (Landy, Maloney & Young, 1990). We asked observers to intercept balls on an approach to their face whilst wearing either telestereoscopes or Cyclopean glasses. The former increase disparity, the latter remove disparity. By running blocks of trials with each device back-to-back we hoped to create circumstances conducive to a reduction in the influence of disparity. No evidence was found to suggest that such an adaptation occurred.

3.2.2 Introduction

Time to contact (TTC) of an object approaching the face is given by changing absolute optic disparity ($TTC_{Face} = \phi/(d\phi/dt)$ and also changing retinal size ($TTC_{Face} = \theta/(d\theta/dt)$). It has been shown that perception of TTC can be influenced by both changing size (looming) and changing disparity (Heuer, 1993; Gray & Regan, 1997; Wann & Rushton, 1995). Judge & Bradford (1988) demonstrated that an observer's ability to catch a ball thrown towards them is disturbed by telestereoscopes (which increase inter-ocular separation and thus also increase disparity, ϕ).

Perception of depth relies on a combination of depth cues (see for example Landy, Maloney, Johnston & Young, 1995). Landy, Maloney & Young (1990) reported that when a cue to depth is unreliable it is given less 'weight', i.e. it has less influence on the final percept of depth. Depth is believed to rely on "weak fusion", this is the combination of estimates from a number of discrete systems attuned to different cues or information sources. The *dipole* model proposed in the previous chapter suggests strong fusion with the early combination of disparity and size, and in its current form does not allow for a change of "cue-weightings".

We wished to determine if it was possible to reduce the 'weighting' on disparity in the perception of TTC. Judge & Bradford (1988) had observers catch balls thrown to them whilst they were wearing telestereoscopes. We took a similar task and modified it to alternate intercepting an approaching ball whilst wearing telestereoscopes, with interception whilst wearing Cyclopean spectacles (which reduce the inter-ocular separation to zero) and normal viewing. Because observers know whether they have successfully intercepted the ball, the task provides unambiguous and clear feedback about whether or not the interception is correctly timed. We wished to test the hypothesis that after a number of alternations between devices the observer would come to rely less on the 'unreliable' disparity cue.

We also inspected our data carefully, to see if it provided us with insight into mechanisms for perception of TTC in front of the face by examining the direction and magnitude of the errors and after-effects following adaptation.

3.2.2.1 Optical devices

Three optical devices were built, a telestereoscope, Cyclopean spectacles and a pathlength modifier (see figure 1 below). Telestereoscopes approximately double the inter-ocular separation so increasing the disparity of every object in the scene. Cyclopean spectacles effectively place a single eye in the centre of the head and thus remove all binocular disparity. The path-length modifiers were a device built to control for the increase in path length (distance the light must travel) associated with the telestereoscopes. All devices had approximately the same path-length and (narrow) field of view.

3.2.2.2 Perceptual effects of different viewing devices

Let us begin with some theory and consider the likely effects on perception of distance with the devices. Perception of distance and depth results from a combination of cues (see Landy, Maloney, Johnston & Young, 1995; Bruno & Cutting, 1988; Johnston, Cumming & Parker, 1993). For simplicity, in the following I will assume a simple linear combination of cues.

In a telestereoscope the disparity of every object is increased. Therefore objects should appear closer to the observer. This fits with the observation by Helmholtz (1910) that a visual scene observed through telestereoscopes appears shrunken, like a scale model.

Cyclopean glasses should have a similar but reciprocal effect. Disparity is reduced to zero. Any element that is primarily specified by disparity should be seen to be more distant. Elements that have their depth well-specified by other cues to distance should be less influenced by Cyclopean glasses. The rationale for this is that disparity only specifies distance when it is not zero. Zero disparity can result from viewing an object at any distance over about 10m away.

The path-length modifiers place the observer slightly back from the scene, therefore the scene should appear slightly smaller.

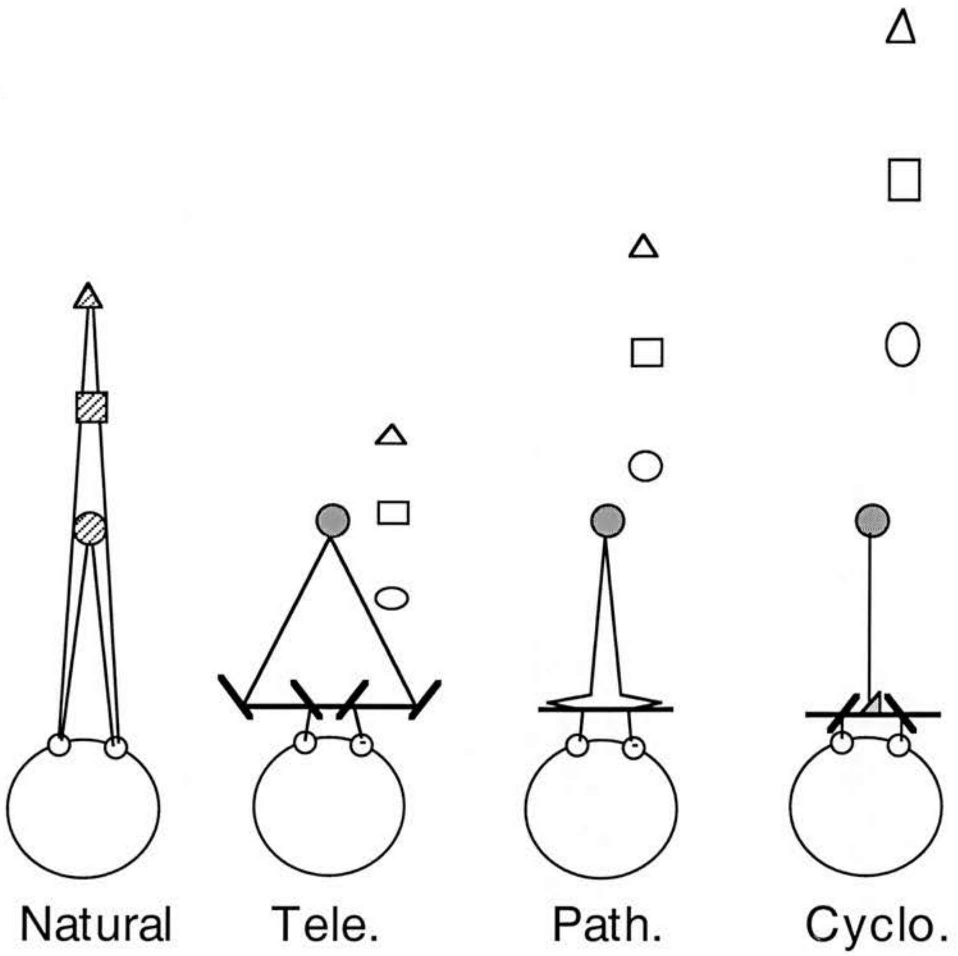


Figure 1: Optical devices. Left panel shows viewing of three objects without an optical devices. Three rightmost show the perceived depth of same objects with optical devices. Depth is calculated on the basis of a linear cue-combination. Empty shapes indicate perceived positions of objects.

3.2.2.3 Experiment, feedback and the manipulation of disparity

Judge & Bradford (1988) threw balls to observers who attempted to catch them one-handed. They assessed performance, adaptation and after-effects of adaptation. Judge & Bradford noted that in their experiment, it was not possible to distinguish spatial errors (reaching to the wrong place) from temporal errors (reaching at the wrong time) in catching.

We designed a variant of their task to restrict errors to timing. The task was modified so that the observer swiped across at a fixed distance to deflect an approaching ball. The hand remained unseen, off to one side and could only be seen very transiently during ball deflection. This arrangement, with a fixed distance and unseen hand was chosen to remove spatial errors: the position of the hand was given by proprioception (felt position of the hand) and was not obviously influenced by the optical device worn. The interception point remained approximately fixed. Therefore, the observer should be able to correctly perceive the position of their hand and errors should be due to wrongly anticipating the arrival of the ball.

The speed of the approaching ball was varied because Judge & Bradford pointed out that if the hand position is fixed, then for a constant ball speed the observer can correct their temporal errors by simply adding an offset onto their normal swipe time.

Consider first the Cyclopean glasses. Disparity is removed, however looming remains unchanged. If an observer makes errors initially, they can reduce errors by taking more note of looming and less of disparity in later trials. With the telestereoscopes the same applies, looming remains unchanged and veridical.

With either device, the simplest way to reduce errors would be to ignore disparity. When blocks of trials with each device are run back to back then ignoring disparity is the only way to avoid errors every time the viewing device is changed.

As noted above, it has been suggested that if a cue is unreliable then the visual system may reduce its "weighting". By running blocks of trials with different optic devices we made disparity very unreliable. Thus we expected to find a reduction in the weighting applied to disparity after repeated trials.

3.2.3 Methods

3.2.3.1 Observers

Five individuals with normal or corrected-to-normal vision participated. They were staff or students of Edinburgh University. All except one observer (John Wann) were naive as to the experimental hypotheses.

3.2.3.2 Task

Observers sat on a chair looking through an optical device. The observer's hand was placed approximately 30cm in front of them off to one side. The hand was not visible. A tennis ball mounted on a rigid pendulum swung towards the observer. The observer had to swipe across in front of their face to deflect the ball. If the ball was missed then the pendulum clattered into a bar placed before the observer's face to prevent the ball hitting them. (see fig 2 below)

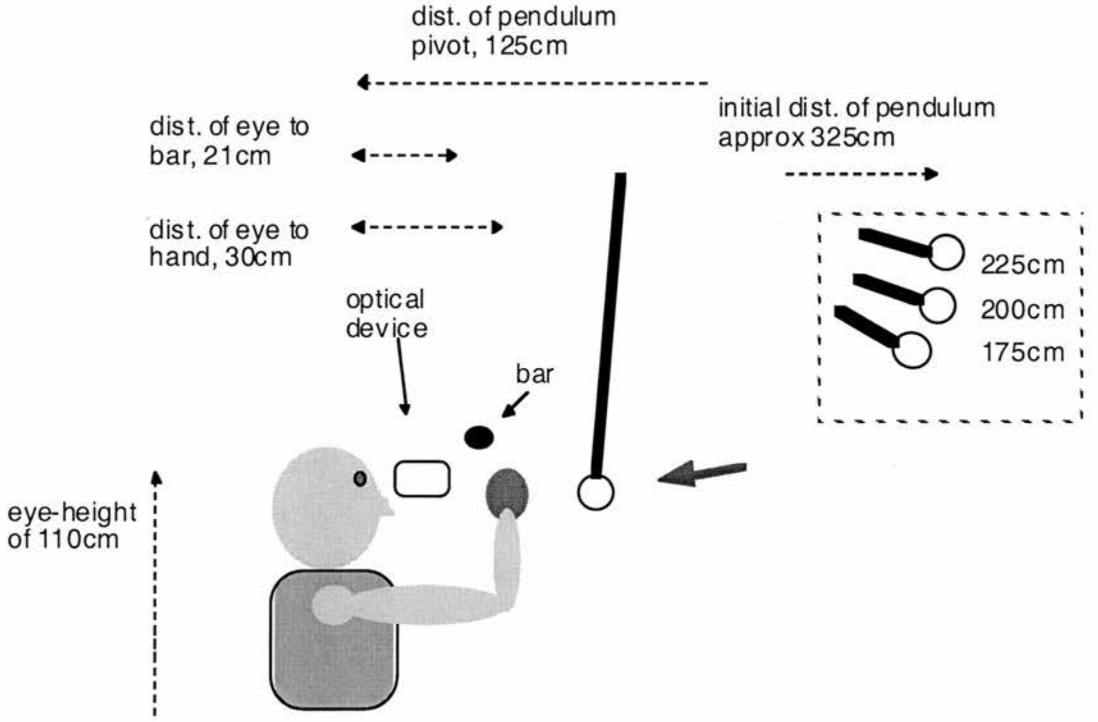


Figure 2: Observer viewing approaching object through optical device. The ball was stopped from hitting the observer by a bar that caught the pendulum.

The ball did not become visible until it was almost at the pivot point (the lowest vertical position) of the pendulum. A window and a variety of objects were visible behind the pendulum and the experiment was conducted in a well lit room. The speed of the approaching ball was varied by releasing the pendulum from 3 different heights (producing speeds in the range of 4-6 m/s). The observer could not see the height of release of the pendulum.

The task provides the observer with naturalistic feedback on each response (contacting, just contacting or not contacting the ball). Every timed swipe recorded is influenced by the current optical arrangement and previous responses.

In the following sections, data analysis is based upon mean response during a block of nine consecutive trials. During the trials, error reduction is occurring. However, Judge & Bradford's data suggests that it takes longer than this to fully adapt to a new optical device. Thus, we can be confident that the results reliably reflect the influence of the optical devices. However, because this error-correction process runs through the data, it is not possible to exactly quantify the pure shift in timing due to a given optical device. This would require a further experiment that did not use feedback.

3.2.3.4 Measures

A Selspot infra-red optical tracker recorded the position of the ball and the hand. The positions were recorded at 312Hz. Time course records were examined to determine (i) the place at which the trajectory of the ball and the trajectory of the hand crossed: the 'interception point', (ii) the time at which the ball crossed the interception point, (iii) the time at which the hand crossed the interception point, (iv) ball velocity. From this it was possible to determine the temporal error in the hand reaching the interception point for each swipe. By comparing this time across conditions it is possible to determine the change in swipe timing associated with the different optical devices.

3.2.3.5 Procedure

Each observer completed the blocks of trials of each condition in the following order. Each block consisted of 9 trials:

Optical Device	Abbreviation
Normal	N
Pathlength	P11
	P12
Telestereo	T1
Cyclopean	C11
	C12
	C13
Pathlength	P2
Cyclopean	C21
	C22
	C23
Telestereo	T21
	T22
	T23

Each block followed straight on from the previous one. Observers closed their eyes whilst the optical device was replaced.

3.2.3 Results

Results from all five observers are averaged (individual data is very noisy) and plotted as shifts in swipe time, compared with normal vision (fig 3). There are two graphs in figure 3 below. The upper graph shows the change in the initiation of the swipe (the time at which the hand starts to move). The lower graph shows the change in interception time. If the hand does not contact the ball then the interception time is the time that the hand crosses the ball's path.

It can be seen from fig 3 that observers attempt to correct their timing errors. The data is plotted using box plots which clearly illustrate average, distribution, skew and outliers (see a standard statistics textbook such as Howell, 1989 for a full description).

Box Plots

Box plots give a ready visual representation of central tendency and distribution of data points in a dataset. The skew and outliers are easily identified. The components that make up a box plot are as follows: The central bar indicates the median. The upper and low limits of the box are known as *hinges*. The *hinges* indicate the upper and lower quartiles, the distance between the two obviously giving the interquartile range which is sometimes known as the *H-spread* when describing box-plots . The bars that bracket the box are known as *adjacentvalues* and are calculated as follows:

lower fence = lower hinge - 1.5(H-spread) [= lower quartile - 1.5(interquartile range)]
upper fence = higher hinge + 1.5(H-spread)

lower adjacent value = smallest value >= lower fence
upper adjacent value = largest value <= upper fence

So the adjacent values are given by the data points that are closest to, but within the fences. Any points which fall outside the fences are *outliers* and are individually plotted.

Worked examples of how box plots are constructed and a more qualified appraisal of their utility can be found in any basic statistical text, a good starting place is Howell (1989).

The central line within the box is the median, the ends of the box are the upper and lower quartile, outliers are individually plotted. When an outlier is found, it is very difficult to determine if it is noise, an erroneous response, or rather reflects the observer's error reduction strategy. The average results are fairly representative of the individual results -none of the individual results were notably different to the mean results displayed below.

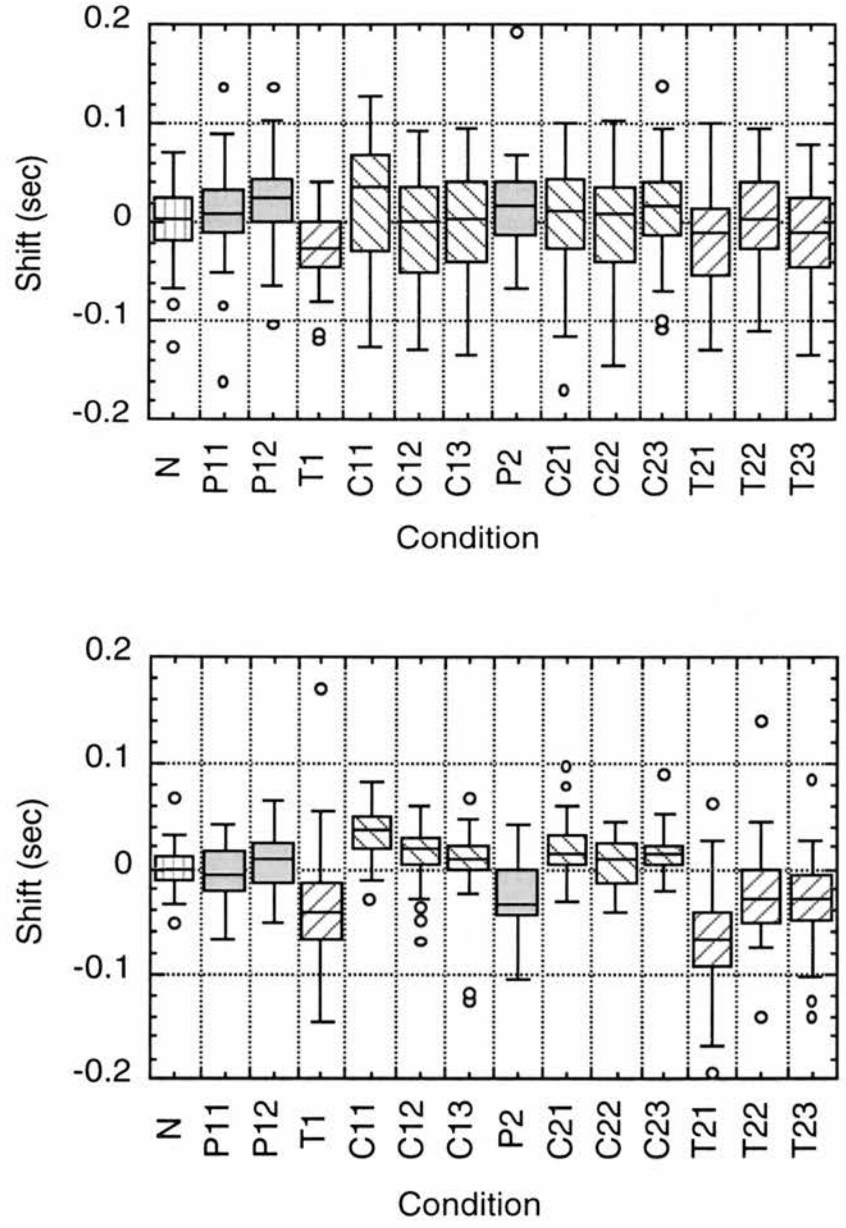


Figure 3: Box plots of swipe times. **Upper.** Mean shift in hand start time (against baseline of normal vision). Conditions in presentation order. Data from all five observers. **Lower.** Mean shift in hand interception time (against a baseline of normal vision).

3.2.3.1 General Observations

The shifts in initiation of swipe and interception are similar. The variance of the interception time is considerably smaller than that for swipe initiation time, therefore interception times will be referred to in the following.

It can be seen that measurable shifts in interception time occur when the optical device through which the observer views is changed. The Cyclopean effect (late swipe) can be seen by comparing C21 against P2, where a negative shift (early swipe) changes quickly to a positive swipe. The telestereoscope effect (early swipe) can be seen by comparing T1 against P12 or T21 against C23. It is evident that observers reduce their temporal errors with repeated trials. For example, from C11 to C13 the positive shift (late swipe) is reduced. Observers are beginning to adapt to Cyclopean glasses. Similarly, notice that from T21-T23 the negative shift (early swipe) is reduced from block to block, showing adaptation to the telestereoscopes. After-effects can also be clearly seen (see T1 vs. T21 or P11 vs. P2). For example, there is a larger negative shift at T21 than at T1. The reason for this is that T21 involves a switch from Cyclopean glasses to telestereoscopes - hence the shift represents both the expected shift due to wearing telestereoscopes and a shift in the same direction due to removal of the Cyclopean glasses. A 'pure' after-effect can be seen when comparing P2 (removal of Cyclopean glasses) with P12 (the 'baseline' swipe time for pathlength modifiers).

3.2.3.2 Switching away from disparity?

During the experiment observers use all three viewing devices. The difference between them is the disparity information that they provide. If it is possible to learn to ignore disparity then it would be expected to be noticeable in the later trials. Learning to ignore disparity would be evidenced by smaller shifts when switching between devices in later trials. Casual inspection of fig 3 indicates no obvious reduction in the influence of disparity. For example, T21 is the last change of optical device but produces the largest shift in grasp timing of all the blocks.

On the other hand, if there was no reduction in the influence of disparity, it would be hypothesised that the T21 shift should be larger than the T1 shift. This is because T21 follows C23, Cyclopean glasses to telestereoscopes is the maximal change in disparity (from zero disparity to exaggerated disparity). Such a difference is found, the shift associated with T21 is significantly larger than T1 ($p < 0.01$).

Similarly if there is no reduction in the influence of disparity then P2 should show a larger shift than P12. We can take P12 as providing a baseline for swipe time with the path-length modifiers. P2 follows C13 and if disparity is not ignored then there should be an after-effect with the P2 shift larger than the P12 shift. This is the case, P2 is found to be significantly different to P12 ($p<0.01$).

3.2.4 Conclusion

The alternation of optical devices dramatically alters the disparity information available. This should make disparity an "unreliable" cue and prompt a change in the use of disparity. However, there is no indication of a reduction in the influence of disparity during the course of the experiment. These results are compatible with a strong model of fusion that does not distinguish between disparity and size and so provides no mechanism by which their respective "gains" may be changed.

See the appendix for discussion of models of perception of TTC with the hand and possible mechanisms for adaptation.

3.3 Appendix

Why did the participants mis-time interception when wearing telestereoscopes and Cyclopean glasses? There are two routes for estimation of TTC_{Hand}. TTC_{Hand} could be estimated directly or TTC_{Hand} could be estimated via TTC_{Face}. Both routes are considered below.

Our data involved feedback and adaptation, so it is not possible to make precise predictions to fit models. However, any plausible model must be compatible with the trends in the data. Below I offer several possible models.

A1 TTC_{Hand}

A1.1 Judge & Bradford's scaling model

This model was proposed (and dismissed by Judge & Bradford, 1988) to account for their own ball-catch results. Judge & Bradford had observer's catch a ball whilst wearing telestereoscopes.

Helmholtz (1910) noted that when viewing the world through telestereoscopes, the world appeared smaller. Therefore, if the observer reaches to the perceived position of the ball, then they will reach short as the world is perceived as smaller (if it is assumed that scaling occurs around the head-position - see fig A1). This prediction was noted to be roughly compatible with the results from the initial adaptation to the telestereoscope. However, it was concluded that the theory was incompatible with the after-effects observed upon removal of the telestereoscopes.

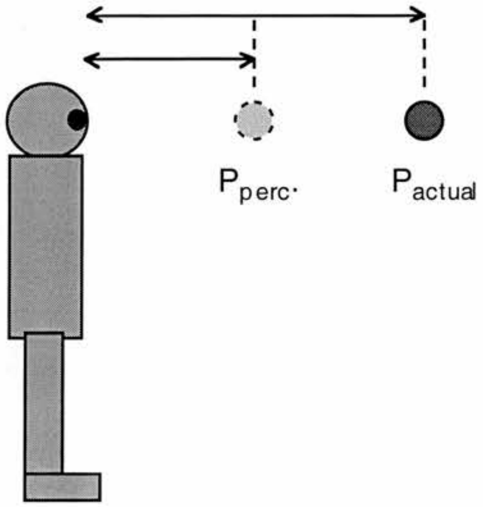


Figure A1. Mis-perception of ball position resulting from viewing with telestereoscopes. Observers reaching to the perceived position ($P_{perc\cdot}$) rather than the actual position (P_{actual}) will reach early, which is compatible with results reported by Judge & Bradford (1988).

A1.2 Judge & Bradford's mis-perception of ball speed

The alternative explanation suggested by Judge & Bradford (1988) relies upon a mis-perception of ball velocity. In their study, they assumed that the distance of the ball is correctly perceived and that ball velocity is mis-perceived. With the telestereoscopes the mis-perception of ball velocity is hypothesised to result from the faster retinal speed due to the increased disparity gradient in the telestereoscopes. Therefore, the ball should appear to be moving faster than it actually is (see fig A2).

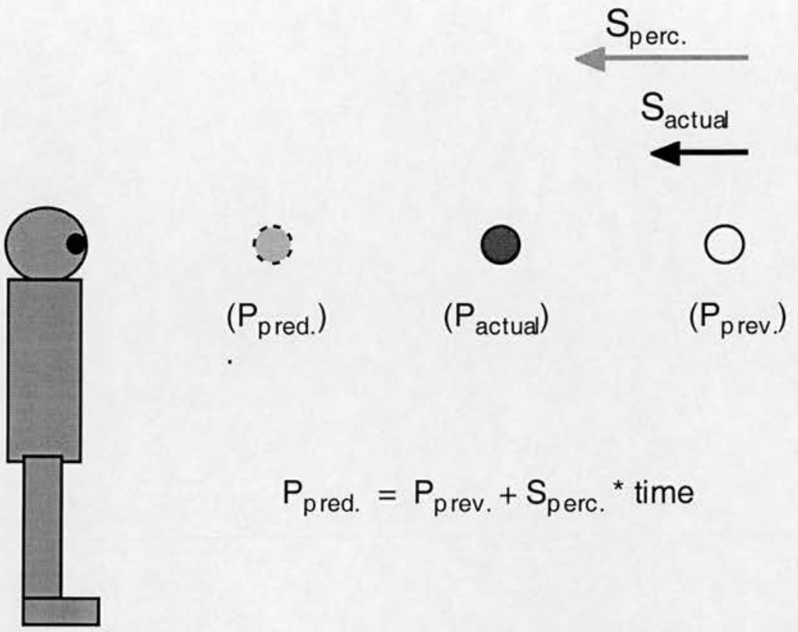


Figure A2. The observer perceives the distance of the ball correctly. Prospective control requires the observer to predict the location of the ball. If the observer misperceives the speed of the ball then the predicted position of the ball will be incorrect. In the above case the ball is perceived to be travelling faster than ball it actually is so the ball will be predicted to be closer (P_{pred}) than it is (P_{actual}) and the observer will reach early. This is compatible with the results reported.

Thus it appears that the ball will reach the observer earlier than it actually does. Therefore, the observer reaches early. The problems with this theory are (i) it relies on the distance of the ball being perceived correctly which is contrary to Helmholtz (1910); (ii) it assumes that an increase in retinal velocity is perceived as an increase in speed of motion-in-depth despite there being no simple relationship between the two (a close object moving slowly produces the same retinal velocity as a distant object moving fast); (iii) it assumes that TTC is perceived from distance and velocity which is contrary to most modern research on TTC (e.g. Lee, 1976; Gray & Regan, 1997). Therefore I do not consider this model to give a satisfactory explanation of our results.

There are two more simple models which I will quickly cover before considering estimations of TTC with the hand (TTC_{Hand}) based upon modifications of estimates of TTC with the face (TTC_{Face}).

A1.3 TTC to a given vergence angle.

The observer could base their catch or swipe on the arrival of the ball at a given vergence angle or absolute optic disparity. Let us consider a simplified case: We will ignore the time it takes to perceive arrival and move a hand and assume that the observer can react and move their hand instantaneously. If the observer wishes to intercept the ball at 50cm from the face then when the ball reaches the disparity equivalent to 50cm from the face, then they could reach their hand across to the point 50cm from the front of their face, and grasp the ball. With the telestereoscopes, the ball would reach a given absolute disparity or vergence angle earlier than normal and so predict the telestereoscope results. However this model falls apart when trying to explain the Cyclopean glasses results, because a vergence angle of zero does not specify any definite distance. I do not consider this model to provide a satisfactory explanation.

A1.4 Distance and speed

TTC could obviously be calculated on the basis of distance and speed.

$$TTC_{Hand} = (D_{Ball} - D_{Hand}) / S_{Ball} \qquad [A1]$$

$$= (D_{Ball} / S_{ball}) - (D_{Hand} / S_{Ball}) \qquad [A2]$$

The scaling of D_{Ball} and S_{Ball} by the optical devices should be isotropic, therefore errors cancel out in D_{Ball} / S_{ball}. However, D_{Hand} is not scaled as it is not in view. Therefore the error should be:

$$TTC_{Error} = -(D_{Hand} / S_{Ball}) \qquad [A3]$$

Here, D_{Hand} is correct (not affected by scaling as unseen). Error comes from S_{Ball} being scaled. If we consider the effect of scaling: with the telesterescopes, the speed will be perceived as being slower and so the observer should swipe earlier. With the Cyclopean glasses, the speed will be perceived as faster and so the observer should swipe later. This is compatible with the results.

A2 TTC_{Face}

There is an alternative to the direct estimation of TTC_{Hand} via the routes above. TTC_{Hand} can be based upon estimation of TTC_{Face} and a "correction".

An estimate of TTC_{Face} could be used during the initial stages to cue up an action. The correction could come into play later as the ball gets closer and the potential to determine distance of the ball or speed of the ball improves. There are a lot of ways that TTC_{Hand} could be estimated via TTC_{Face}. Below I cover some of them. First however it is necessary to look at different routes to estimation of TTC_{Face}.

A2.1 TTC_{Face}: *Distance and velocity*

The most intuitive way to calculate TTC is on the basis of ball distance and speed.

$$TTC_{Face} = D_{Ball} / S_{Ball} \qquad [A4]$$

If the scaling of ball distance and ball speed is isotropic with the optical devices then there should be no error in estimation of TTC_{Face} when wearing either telestereoscopes or Cyclopean glasses.

A2.2 $\boldsymbol{TTC_{Face}}$: **Dipole** *model*

This model which was proposed in the previous chapter combines disparity and retinal size early and has a common motion-in-depth mechanism in line with that proposed by Regan & Beverley (1979). The algebraic formulation of the dipole model is:

$$TTC_{Dipole} = (\phi + \theta) / (d\phi/dt + d\theta/dt) \qquad [A5]$$

Telestereoscopes:

$$TTC_{Dipole} = (2\phi + \theta) / (d(2\phi)/dt + d\theta/dt) \qquad [A6]$$

The changes in θ resulting from increasing interpupillary distance are cancelled by similar changes in $d\theta/dt$.

Cyclopean glasses: $\phi = 0$, $d(\phi)/dt = 0$

$$\therefore TTC_{Dipole} = (\theta) / (d\theta/dt) \qquad [A7]$$

Therefore, none of the optical devices should change estimate of TTC_{Face} with the *dipole* model.

A2.3 $\boldsymbol{TTC_{Face}}$: *Looming (tau)*

$$TTC_{Face} = \theta / (d\,\theta/dt) \qquad [A8]$$

None of the optical devices change retinal size (θ) and rate of change of retinal size (looming, $d\theta / dt$). Therefore, if perception of TTC_{Face} is derived solely from optical size, then it should be unchanged by any of the optical devices.

A2.4 $\textbf{\textit{TTC}}_{\textbf{\textit{Face}}}$*: Weak Fusion*

TTC_{Face} could be calculated by obtaining independent estimates from disparity (TTC_{Disp}) and size (TTC_{Size}) and then combining them.

$$TTC_{Face} = aTTC_{Disp} + bTTC_{Size} \qquad [A9]$$

Such a model would correctly estimate TTC_{Face} with the telestereoscopes (changes in θ resulting from increasing inter-ocular separation are cancelled by similar changes in $d\theta/dt$).

However, any simple additive (or multiplicative model) would over-estimate TTC_{Face} with cyclopean glasses. This is because TTC_{Disp} would remain at infinity. This makes the cyclopean glasses a very good device for strongly testing models of TTC. We did not directly test this hypothesis - remember our experiment required estimate of TTC_{Hand}.

A2.5 $\textbf{\textit{TTC}}_{\textbf{\textit{Face}}}$ *via looming, distance and IPD*

Gray & Regan (1997) provide the following

$$TTC_{Face} = I / [\, Z\, (d\phi/dt)\,] \qquad [A10]$$

where I = assumed inter-pupillary distance, Z = distance, ϕ = absolute disparity or vergence angle.

Z can be removed by substituting an estimate of Z based upon retinal size and known size of the approaching object. It is reasonable to assume that the observer has access to known size as they see the same ball every time.

Using a small angle approximation

$$Z = R / \theta \qquad [A11]$$

where θ = retinal size.

so substituting [A11] into [A10] gives an alternative form involving known size.

$$TTC = I\theta / R(d\phi/dt) \qquad [A12]$$

Mis-perception of I (believing that your inter-pupillary distance is unchanged) will lead to systematic errors in perception of TTC and in the case of telestereoscopes lead to the observer estimating that TTC is less than it actually is which is in line with the results. However, this model runs into problems with Cyclopean glasses. Again, $d\phi/dt$ is not specified and the model cannot consider what happens with the Cyclopean glasses.

A3 TTC_{Hand} via TTC_{Face}

So how do we get from TTC_{Face} to TTC_{Hand}? Although this problem is by no means trivial, it is rather regularly side-stepped, most models of TTC are models of TTC_{Face}.

TTC_{Hand} is algebraically given as:

$$TTC_{Hand} = TTC_{Face} - D_{Hand} / S_{Ball} \qquad [A13]$$

Where TTC_{Hand} is the TTC with the hand, D_{Hand} is the distance of the hand from the face and S_{Ball} is the speed the ball is travelling towards the face (see fig A3).

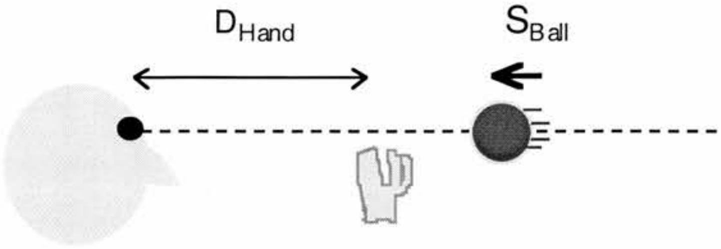


Figure A3: TTC is calculated with reference to the face. The observer works out the TTC_{Face} value at which the observer needs to reach to intercept the ball (see [A13]).

If the hand is in view the necessary correction (and indeed TTC_{Hand} directly) is optically specified. The correction can be determined by use of relative disparity (the difference in optic or retinal disparity between the hand and the approaching object). Alternatively, if the ball approaches from an eccentric direction directly towards the position of the hand then lateral gap closure can be used.

However, when the hand is not in view, then there is no optical specification of TTC_{Hand}, but humans do appear to be able to intercept approaching objects when they do not have sight of a hand. Indeed a cricket batsman is required to intercept the approaching ball with an unsighted bat and this can evidently be done with great accuracy. Therefore, TTC_{Hand} must be able to use intended hand position (the place you expect to intercept the object).

Returning to the above equation [A13], although this is an algebraic specification of TTC_{Hand} it is not clear that this is the route used by the visual system. However we can start with this formulation.

For the purposes of the following we will assume that TTC_{Face} is correctly perceived. this could be achieved by any of the models described in A2.1, 2.2, 2.3, but not 2.4 or 2.5. Further, because the hand is unsighted and so unchanged by the different optical devices. D_{Hand} is given by proprioception (felt position) and so there should be no error in D_{hand}[2]. This assumption is also used in the following.

A3.1 $\boldsymbol{TTC_{Face}}$ *+ mis-perception of ball speed, variant 1*

I noted above that from a linear cue combination it should follow that objects are seen closer with telesterescopes (a "scale model") and further away with cyclopean glasses (see fig 1). If it is assumed that scaling of distance and speed is isotropic then the ball should be seen to be travelling more slowly when wearing the telestereoscopes (spatiotopic speed is spatiotopic speed covered in a unit of time) and faster when wearing the Cyclopean glasses.

Therefore, if TTC_{Face} is correct then when wearing telestereoscopes the observer should swipe earlier and when wearing the Cyclopean glasses swipe later (see equation [A13]). This is compatible with the results.

A3.2 $\boldsymbol{TTC_{Face}}$ *& mis-perception of ball speed, variant 2*

This relies on the same misperception of ball velocity as in A1.2 (faster retinal speed = faster spatiotopic speed). However, the estimation of TTC_{Face} is by a different route.

$$TTC_{Hand} = TTC_{Face} - D_{Hand} / S_{Ball} \qquad [A14]$$

If TTC_{Face} is estimated correctly, then with telestereoscopes, S_{Ball} should be mis-perceived as higher than it really is. This should thus reduce the time

---

[2] It is possible that perceived distance of the hand is changed. It is known that wearing field reducing glasses (leaving approx 12 degrees field of view) leads to the perception that eye-height has changed (see Dolezal, 1982). This is an exproprioceptive (the position of the body relative to the environment) change. It may be possible that some visual manipulations can

at which the observer reaches. The swipe time will be smaller than the actual value and so the observer will reach too late. This is clearly incompatible with the results reported here.

A3.3 $\boldsymbol{TTC_{Face}}$ *& the "Wann model" of* $\boldsymbol{TTC_{Hand}}$

The model that remains the favourite of John Wann. It uses D_{Ball} instead of $S_{ball.}$

$$TTC_{Hand} = TTC_{Face}\,(1 - D_{Hand} / D_{Ball}) \qquad [A15]$$

Where D_{Hand} is the distance of the hand and D_{ball} is the distance of the ball. Both are measured with the observer's face as the origin. With the telesterescopes D_{ball} will be perceived as closer. Critically when $D_{ball} = D_{Hand}$, D_{ball} will be perceived as closer than D_{Hand} . Therefore, if TTC_{Face} is perceived correctly, then the observer will swipe early. With the Cyclopean glasses the D_{ball} will be perceived as more distant, therefore the observer will swipe later. These predictions are compatible with the results.

A4 Adaptation and after-effects

Temporal errors are slowly reduced. The results of Judge & Bradford show a slow downward descent to the correct TTC swipe time. This is compatible with the data reported here. What scope is there for adaptation, where might it be occurring? I will not consider all the candidate models of TTC_{Hand}. Rather, I will concentrate on the estimation of TTC_{Hand} via TTC_{Face}. Figure A4 below indicates some potential sites for adaptation.

---

lead to a change in proprioception. However, I'm trying to keep things simple, so I keep to the assumption above.

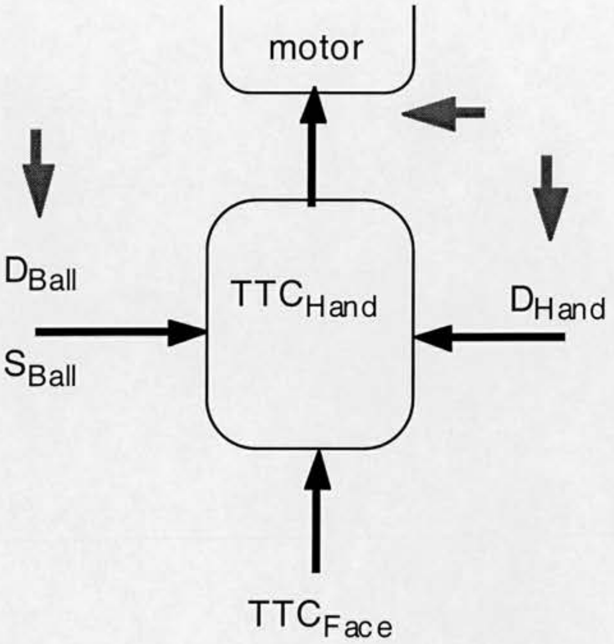


Figure A4: Potential sites of adaptation (indicated by red arrows).

A4.1 Adaptation in perception of D_{Hand} and S_{Ball}

In the experiment reported, D_{Hand}, is seen transiently during each swipe so there is some scope for an adaptation. There are many examples of adaptation of 'felt position', for example when wearing prisms (Rock, 1966).

It may be possible to modify S_{Ball}. It is not clear how this change could occur, but the same functional change can be achieved by scaling D_{Hand} by x or S_{Ball} by $1/x$. Let us assume that S_{Ball} is the seat of adaptation so that we can work through some predictions.

As noted above, when wearing the Cyclopean glasses observers should perceive the ball as moving faster than it is and so swipe late. Error correction should lead to a downwards revision of S_{Ball}. Therefore if the observer then switches to a different device then they will under-estimate S_{Ball} and consequently swipe early.

When wearing telestereoscopes the observer should perceive the ball as moving more slowly and so swipe early. Error correction should lead to an upwards revision of S_{Ball}. Therefore if the observer switches to a different device then they will over-estimate S_{Ball} and consequently swipe late.

A4.2 Adaptation at the output stage

A simple adaptive strategy is to modulate the output of the perceptual system, or the firing threshold of the motor system.

Judge & Bradford (1988) pointed out that if the position of the hand was fixed then a correction could be achieved through:

$$TTC_{Swipe(new)} = TTC_{Swipe(old)} + a \qquad [A16]$$

This would be easily achieved by just initiating the motor action at a different trigger time.

When there is a range of ball speeds then the simple addition of an offset will not work. However the following should work.

$$TTC_{Swipe(new)} = TTC_{Swipe(old)} * a \qquad [A17]$$

This would require that the output of the perceptual system is modulated (the gain reduced) or the sensitivity of the motor system decreased (so higher trigger values are necessary).

Such a modification as above [A17] would allow for errors to be reduced. If we consider the after-effects then with the telestereoscopes it is necessary to reduce the gain of the output of the perceptual system. When moving to a new device this would predict that the observer would swipe later than normal. With the Cyclopean glasses the opposite would occur.

A4.3 Compatibility with results

After-effects are compatible with the predictions arising from a modification of S_{Ball}, D_{Ball}, D_{Hand} or at the percepuo-motor junction. For instance the telestereoscope after-effect can be seen by comparing C11 against C21, the Cyclopean after-effect can be seen by comparing P2 and P12. It is not possible to use the data from this chapter to distinguish between these possibilities

The visual system does not have access to a system's diagram of the estimation of TTC_{Hand}. The visual system will not be able to identify where the error is arising nor know how to optimally deal with it. Given just an error signal then the adaptation that occurs should simply reflect the plasticity and time-course of adaptation of the various components within the system.

A5 The problem of S_{Ball}

A change in retinal velocity or retinal size does not specify a spatiotopic speed. This is because a far object moving fast or a close object moving slowly can produce the same change in disparity. It is possible to determine ball speed from a function of the rate of change of disparity and the rate of size.

Alternatively known size of the ball can be used in conjunction with retinal size to obtain an instantaneous estimate of distance. Lastly the next temporal derivative of changing size or disparity could be used to disambiguate between a fast distant object and a slow near object if constant speed is assumed.

Regan has shown that even when there is no information as to spatiotopic speed a percept of spatiotopic speed still arises: he had observers view monocularly a looming square (sinusoidal modulation of size) and they reported a definite percept of spatiotopic speed. This suggests that they must have relied on a default spatiotopic speed or assumption of viewing distance.

3.4 References

Bruno, N. & Cutting J.E. (1988). Minimodularity and perception of layout. *Journal of Experimental Psychology: General*, 117, 161-170.

Gray R. & Regan, D. (1997). Accuracy of estimating time to collision using binocular and monocular information.*Vision Research*, 38, 499;

Helmholtz, H. (1910). *Handbook of Physiological Optic, Volume III: Translation of 3rd edition.* Edited Southall, J.P.C. (1962) Dover Publications, In, New York, USA.

Howell, D.C. (1989). *Fundamental statistics for the behavioral sciences: 2nd edition.* Duxbury Press, Belmont, CA, USA.

Heuer, H. (1993). Estimates of time to contact based on changing size and changing target vergence. *Perception* 22, 549.

Johnston E.B., Cumming B.G., & Parker A.J. (1993). Integration of Depth Modules: Stereopsis and Texture. *Vision Research*, 33, 813-826.

Judge, S.J. & Bradford, C.M. (1988). Adaptation to telescopic viewing measured by one-handed ball-catching performance. *Perception*, **17**, 783-802.

Landy, M.S., Maloney, L.T. & Young, M.J. (1990). Psychophysical estimation of the human depth combination rule. In Schenker, P.S. (Ed.) *Sensor Fusion III: 3D Perception and Recognition, Proceedings of the SPIE*, 1383, 247-254.

Landy, M.S., Maloney, L.T. Johnston, E.B. & Young, M. (1995). Measurement and Modelling of Depth Cue Combination: in Defense of Weak Fusion. *Vision Research*, 35, 389-412.

Rock I. (1966). *The Nature of Perceptual Adaptation.* Basic Books, London.

Wann, J.P. & Rushton S.K. (1995). Grasping the impossible: Stereoscopically presented virtual balls In B. Bardy, R. Bootsma & Y. Guiard. (Eds) *Proc. of the 8th International Conference on Event Perception and Action.* 207-210.

Chapter 4:

Depth and the perception of direction of locomotion

4.1 Foreword

In this chapter I examine the role of depth information. It has been reported that depth aids perception of locomotor direction. This finding is in contradiction to the hypothesis that modular systems use a minimal set of information. An implementation that used depth would either introduce redundancy (if it processed depth internally) or require a feed from a general depth system. In the latter case it seems highly likely that this would slow perception of locomotor direction. Depth is not normally necessary with a rich flow field. In this chapter we examine the use of depth in the active control of steering.

4.2 Paper

From an early draft of: Rushton, S.K., Harris J.P. & Wann, J.P. (1999-in-press). Steering, Optic Flow and the respective importance of depth and retinal motion distribution. *Perception.*

4.2.0 Abstract

Movement through an environment produces an optical spatio-temporal pattern, known as a flow field (Gibson, 1950). When visually guiding movement using a flow field, do humans make use of information about the distance of constituent elements? Employing a novel active steering task, we examined the use of depth (height in scene and disparity) and the role of the retinal motion distribution in the perceptual control of heading from flow. We found that retinal motion distribution, rather than depth order, has the primary role in determining the accuracy of steering.

4.2.1 Introduction

During self-movement, instantaneous heading and path are mathematically specified in the retinal flow field. In principle, knowledge of the depth of objects in a flow field simplifies the determination of heading. In particular, it has been argued that during a pursuit eye-movement depth may assume a critical role in the perception of heading (van den Berg & Brenner, 1994a). If an observer makes an eye-movement during locomotion, then the retinal flow field is complicated (Regan & Beverley, 1982). van den Berg has hypothesised that the specific role of depth is in identifying near and distant objects: the retinal motion of near objects is primarily determined by observer movement (translation), the motion of far objects is primarily determined by eye movement (rotation). Thus, if the depth of respective objects in the scene is known, a perceptual system can utilise the movement of distant objects to obtain the rotational component. Rotation could be subtracted from the overall flow to give the translational component (van den Berg & Brenner, 1994a).

Consensus has been building for the use of depth in the *perceptual judgement* of heading from flow (Cutting, 1996; Palmisano, 1996; Warren, 1995). There is a performance difference associated with locomotion over a ground plane versus through a cloud (van den Berg, 1992). It is hypothesised that this occurs because a ground plane provides depth order - the optical height of an element is related to its distance away. It has also been reported that stereoscopic depth improves performance in the presence of noise (van den Berg & Brenner, 1994b) and that distant points appear to be necessary for accurate heading judgements (van den Berg & Brenner, 1994a). Furthermore, the incorporation of depth-order in recent computational neural models of human heading perception (Warren, 1995) is bolstered by the finding that neurones in brain area MSTd (a candidate for involvement in the perception of locomotion) are sensitive to binocular disparity (Roy & Wurtz, 1990). However, some recent evidence (Ehrlich et al, 1998) suggests that depth may *not* be used in a complex heading judgement task, making this still a somewhat controversial issue.

The role of depth in *steering* (the perceptual control of locomotor heading) has not been studied experimentally. There are two problems with using the findings outlined above in the context of a general model of heading control. First, the advantages associated with using depth have been questioned (Stone & Perrone, 1997). Second, the perceptual control of action may involve different systems to those implicated in perceptual judgements (Gibson, 1979; Goodale & Milner, 1992; Milner & Goodale, 1996; Nakayama, 1994). We constructed two experiments aimed at re-appraising the findings on the use of depth in heading judgements, in the context of heading control.

4.2.1.1 Ground plane vs. Cloud - an issue of depth or distribution ?

The performance difference associated with locomotion over a ground plane as compared to through a cloud (van den Berg, 1992) can be accounted for without recourse to depth-order information. There is a marked difference in the spatial distribution of dots (and motion vectors) in the upper and lower portions of the visual array resulting from a cloud versus a ground plane display (see fig 1). In previous heading judgement studies, a fixed number of dots were distributed over a plane or through a cloud. For a ground plane, all the dots fall in the lower visual field with the highest density just below the point of fixation. For a cloud the distribution is relatively uniform. In addition, the mean retinal dot speed may also differ between the two conditions, unless this is experimentally controlled. Because accuracy of heading perception will, in part, be determined by the sensitivity to the local motion vectors that describe the flow field, it is likely that both of the distribution parameters (locus and mean speed) of retinal motion will influence the precision of heading judgements. A further, more ecologically motivated, difference is the suggestion that surfaces, such as the ground plane, may have a special status in human vision (Gibson, 1979; He & Nakayama, 1994). Experiment 1 was designed to tease apart the role of the retinal motion distribution and 3D environmental layout in the control of heading from flow. We designed a new display, the 'capped-cloud', with an identical retinal motion distribution (locus and mean speed) to a ground plane, but without its height-in-scene depth order. We were thus able to test whether performance differences between a ground plane and 3D cloud display were due to differences in retinal motion distribution or to differences in depth.

4.2.1.2 Binocular advantage - an issue of depth or correspondence?

The second line of evidence in favour of the use of depth-order in heading was that when binocular disparity was added to a display, noise was less disruptive to the judgement of heading (van den Berg & Brenner, 1994b). It maybe be argued, however, that disparity could have improved performance without providing a specific depth cue. An alternative role for disparity could be in solving the motion correspondence problem of matching of dots between frames. Labelling elements with disparity allows a dot to be disambiguated from its neighbours in the Cyclopean array. Hence it is possible that the previously reported disparity advantage could be due to a simplification of the matching problem rather than depth ordering. In Experiments 2 and 3 we explored the contribution of binocular disparity to the control of heading, by comparing performance for displays with veridical disparity (that provides depth and correspondence information), non-veridical disparity (correspondence information only) and no disparity.

4.2.2 Experimental Design and Rationale

Previous studies have explored the perception of linear heading and curvilinear paths using psychophysical judgement tasks. In this paper we are concerned with a different issue - which properties of the visual flow field inform the active perceptual control of locomotor heading (steering). Strategies and features exploited in steering may differ from those used in judgement tasks (Warren, 1995; Nakayama, 1994). Indeed, it has recently been suggested that there may be separate visual streams involved in perceptual discrimination and perceptual control of action (Goodale & Milner, 1992; Milner & Goodale, 1996). We employed a task that closely matched the demands of natural steering with trials that lasted for 8 seconds, to capture steering as the output of a perceptual-motor feedback loop.

These experiments were not intended to address the *strategies* that observers might use to determine heading from flow and so we did not attempt to restrict the use of any potential strategy that could be employed in the natural world. However, a strategy that can be naturally used is to couple gaze and heading - to look where you want to go and then bring heading direction round to coincide (close the gaze angle down to zero). Allowing use of this strategy was not desirable for two reasons. First, this paper is concerned with visual information and flow, therefore we wished to preclude the use of extra-ocular information. Second, in laboratory simulations, when gaze and heading are coupled, the observer can use two strategies based upon artifactual cues: maintaining the edge rate (number of elements passing the edge of the display) at the left and right edges of the screen, or centring the target in the middle of the display. Either of these strategies (which would not be available under natural conditions) will ensure the observer is heading directly towards the target. Therefore we chose to uncouple gaze and heading. It has been reported that the natural preference of the visual system during steering is to keep the object that is guiding movement in central vision (Land & Lee, 1994). The majority of heading judgement studies do not respect this. However, one paradigm that maps onto this observation, is the film-making procedure of 'Dolly & Pan' (Cutting et al, 1992). In this paradigm the display corresponds to an observer (or camera) fixating a target whilst translating to the left or right of it. This fulfils our two requirements to keep the reference object in central vision whilst dissociating gaze and heading. We used an active 'Dolly & Pan' task with simulated locomotion on an initial trajectory randomly to the left or right of a target tower, fixed in the environment. The observer's task was to control heading by steering using a joystick, so as to travel straight to the tower. Gaze was continuously computer adjusted to keep the tower centred on the projected display.

Some heading judgement tasks lock gaze rotation to a fixed near environmental feature (e.g. Warren & Hannon, 1988) whereas some add a constant rotation of gaze (e.g. Royden et al, 1992). The latter corresponds to visually tracking a target that is circling around the observer. Given the relative commonality of the two situations (Warren, 1995) we chose to use a ground fixation.

Lastly, some studies use non-expanding dots and some use sparse environments with a ground plane and a small number of objects. We decided to use the former so as to allow a more direct comparison with the majority of previous studies.

4.2.3 Conditions and Predictions

4.2.3.1 Experiment 1: The role of retinal motion distribution and depth

We created four conditions. A cloud and standard ground plane (balanced for total number of dots and mean 2D speed) allowed comparison of our steering results with previous judgement experiments (figure 1, i-iv).

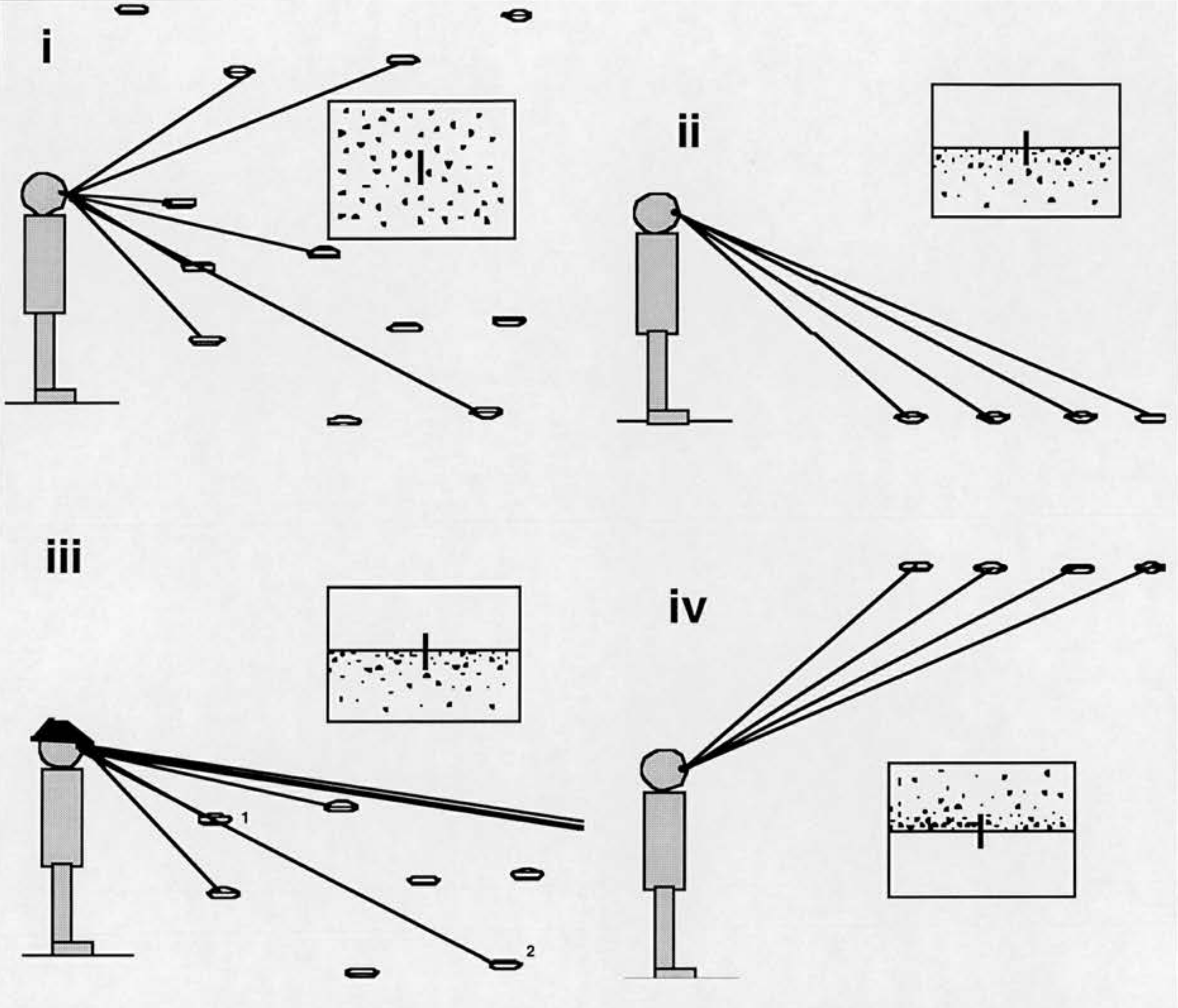


Figure 1 The 4 display configurations were: **(i)** a standard cloud of dots with uniform 3D dot density and dots extending well above and below eye-height; **(ii)** a standard ground plane with uniform density along the plane; **(iii)** a capped cloud; **(iv)** a ceiling (reflection of the ground plane about eye level). Insets show 2D projection arrays.
Dots beyond 40m from the viewpoint were culled from view in all except the capped cloud condition. The capped cloud was a display with a 2D layout (retinal locus and mean speed) identical to the ground plane, but without the height-in-array depth order. To achieve this we took a standard cloud stimulus and applied the following manipulations: (a) increase dot density; (b) reduce the extent of the cloud below the observer's head to just a little (approx. 1m) beneath the observer's feet; (c) extend the culling distance from 40m by approx. 10m; (d) mimic the vision of a flat cap wearer with the peak of the cap tilted slightly downwards. This leaves a display in which the depth of a point in the capped cloud condition is ambiguous. It can be seen in (iii) that points (1) and (2) will appear at the same height in the 2D array. Note a point at the horizon may be at any depth (0.001m to 50m) but a point towards the bottom of the display is restricted to a range of near distances (0.001m to 2.1m).

A 'ceiling' which was a reflection of the ground plane about the horizontal axis (the same depth order information as the ground plane, and also a surface, but much less likely to occur in the real world) which allowed us to examine upper/lower visual field differences. A 'capped cloud' that had an identical retinal motion (mean speed and retinal locus) to a ground plane but lacked its depth order information (see Methods for details). If retinal motion distribution is the critical factor affecting performance, we expect similar performance in the capped cloud and ground plane conditions. If depth is critical we expect performance for the capped cloud to be poorer than for the ground plane.

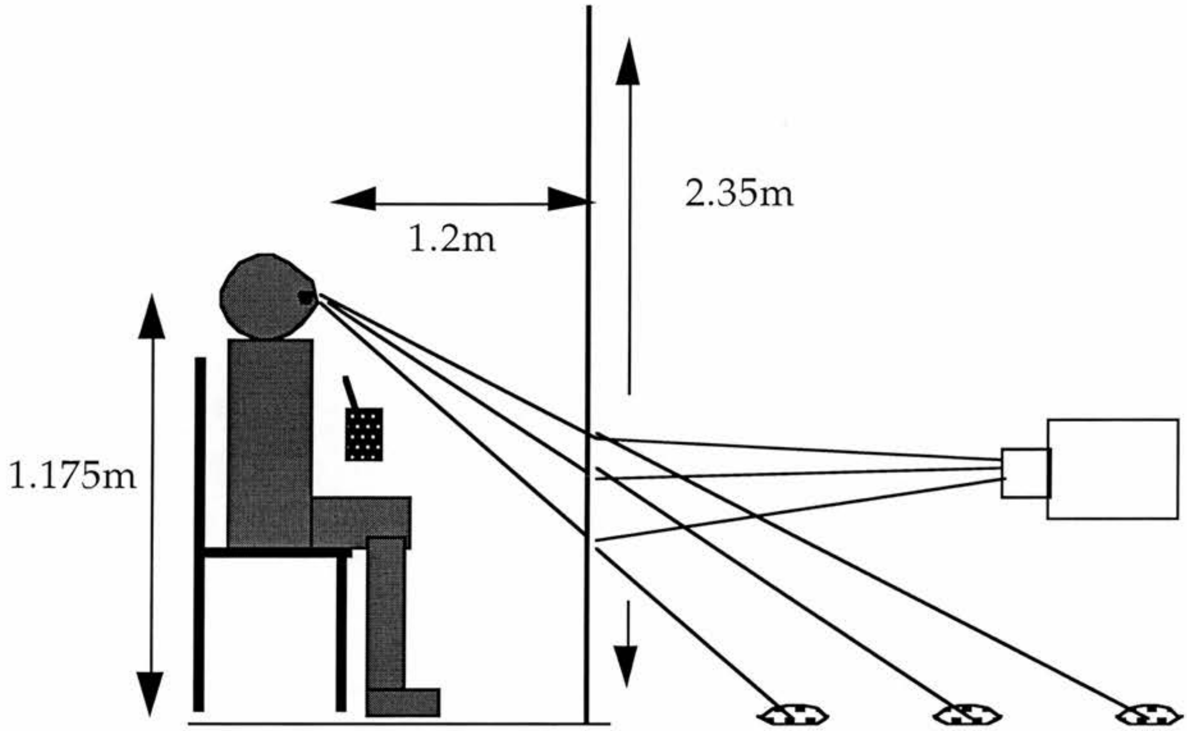


Figure 2 Back projection screen (80^{o} x 80^{o}) with a resolution 1024x768 pixels. The analogue joystick was fed through an 8-bit analogue to digital converter. Latencies incurred in reading the device were less than 1 millisecond so allowing responses to be used in the rendering of the next frame. The joystick controlled the curvature of the heading trajectory rather than the absolute heading angle (similar to a car).

4.2.3.2 Experiment 2: The role of disparity depth

We used a conventional Wheatstone stereoscope to allow the introduction of binocular disparity depth and created three conditions: (i) dots with veridical disparity and disparity defined motion-in-depth; (ii) dots with non-veridical disparity and also disparity defined motion-in-depth that was non-veridical. (iii) dots with no disparity or no disparity defined motion-in-depth, a 'synoptic' display (van den Berg & Brenner, 1994b). In (ii) the dots were assigned disparities and corresponding motion-in-depth trajectories at random (drawn from a distribution with the same ranges as for the veridical case) which did not necessarily correspond to the depth defined by the motion flow.

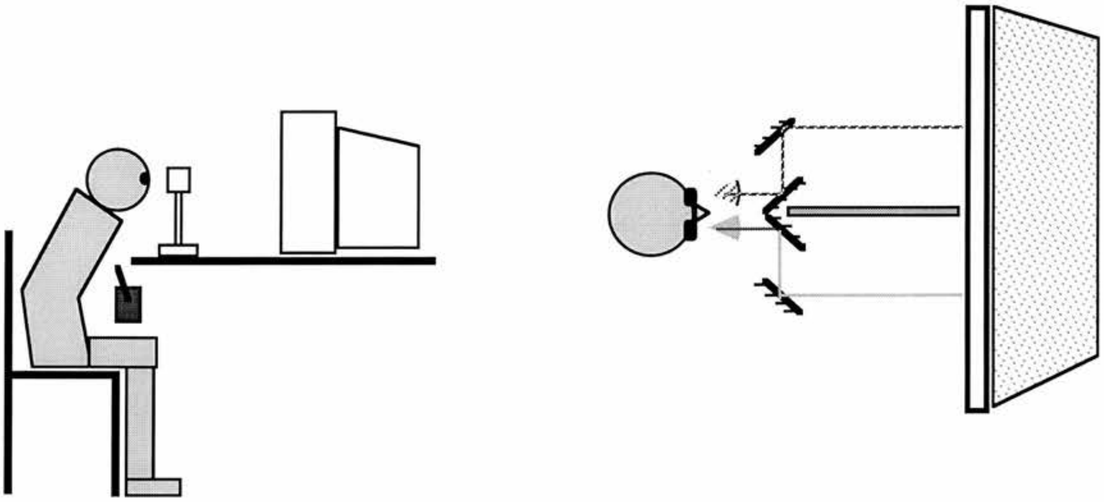


Figure 3. CRT viewed through a Wheatstone stereoscope (horizontal field of view 20°). Display resolution was 481x768 pixels per eye with 2D sub-pixel addressing (32 calibrated linear grey levels) using the 'Backus method' (Georgeson, Freeman & Scott-Samuel, 1996) to increase spatial resolution. Disparity was introduced into the displays by shifting the horizontal position of the dots in the left and right images. For each dot, the magnitude of the pixel shift was calculated based on the distance from the observer of the fronto-parallel plane that contained the dot (using the small-angle approximation, Cormack & Fox, 1985). In the non-veridical disparity condition the pixel shift was calculated using the distance from another point in the scene, for instance the disparity of dot x would be based upon the distance of dot x-1, the disparity of dot x+1 was based upon the distance of dot x. Dots were distributed randomly in space so there was no correlation between the disparity of a dot and its distance (and hence motion in the fronto-parallel plane of the cyclopean array). This procedure ensured that the distribution of disparity depths was identical in the veridical and non-veridical conditions. Also, it produced coherent motion-in-depth trajectories for the dots in both conditions. In Experiment 3, noise was introduced using a similar technique to that previously employed (van den Berg & Brenner, 1994). The 2D velocity for each point was calculated and a noise component calculated and added that had the same magnitude but a random direction. This produced a signal-to-noise ratio of 1 which was the highest employed previously (van den Berg & Brenner, 1994). The same noise was added in all conditions and in the stereo cases the noise was always of the same sign for left and right eyes so producing a perturbation in the xy-plane rather than in z-depth.

If disparity depth order is used in the perceptual control of heading (steering) from flow then performance should be elevated in condition (i) and impaired in condition (ii) relative to the non-stereo conditions. If disparity aids matching then conditions (i) and (ii) should reveal a similar increase in performance.

We do not believe that it should be necessary to add a high level of noise to reveal a use of depth in steering. The only rationale for this would be to counteract ceiling effects with the dependent measures. The results from the first experiment demonstrate the time-course measures we employ are sensitive to small performance differences. Consider a widely accepted example of links between two neural systems, the link from vergence to accommodation, here we find that manipulation of the first system produces measurable changes in the second under natural viewing conditions. For a formal description of the features of 'coupled-systems' see Shallice (1988).

4.2.3.3 Experiment 3: Disparity depth and noise

Experiment 2 should establish whether disparity depth is used in steering. Experiment 3 was the same as Experiment 2 but introduced a high level of noise (SNR=1) to allow comparison of the results with van den Berg & Brenner's perceptual judgement studies.

4.2.4. General Methods

4.2.4.1 Participants

Observers with a range of experience at performing the heading task were deliberately chosen. The spread of relative observer performance in figs 2c, 4a&b reflects the range of experience. With the exception of SKR, observers were naive as to the experimental hypotheses.

4.2.4.2 Apparatus & Displays

The environments were built from randomly distributed dots with a constant number (E1: 400 ±10%; E2&3: 100 ±10%) in view at any time. Each dot had a lifetime of less than 0.5 sec (E1: ~400msec; E2&3: ~440msec). Initial direction of travel, at a speed of 2m/s was to the left or right (E1: 10^{o} ±20%; E2&3: 15^{o} ±20%) of a target tower set 20m away. Temporal resolution was as high as could be consistently sustained (E1: 30Hz; E2&3: 18Hz). Display screens: E1 - Back projection screen (80^{o} x 80^{o}) with a resolution 1024x768 pixels; E2&3 (see fig 2) - CRT viewed through a Wheatstone stereoscope (horizontal field of view 20^o) (see fig 3). Display resolution was 481x768 pixels per eye with 2D sub-pixel addressing (32 calibrated linear grey levels) using the 'Backus method' (Georgeson et al, 1996) to increase spatial resolution. Binocular disparities ranged from approximately 0 to 20 min. The screen (or angular) velocities of individual dots depended on their distance from the observer, their eccentricity, and the course steered by the observer: a large heading error results in additional projected lateral motion and fast dot speeds. As an example, consider a dot positioned at 20 deg eccentricity, 4m from the observer, who is moving at 2m/s (a very close and very eccentric point). Its velocity would be 23 deg/s. Displays were viewed in a darkened room.

4.2.4.3 Performance measures

Heading error (gaze-heading angle) was recorded throughout the trial (see figs 4 & 5).

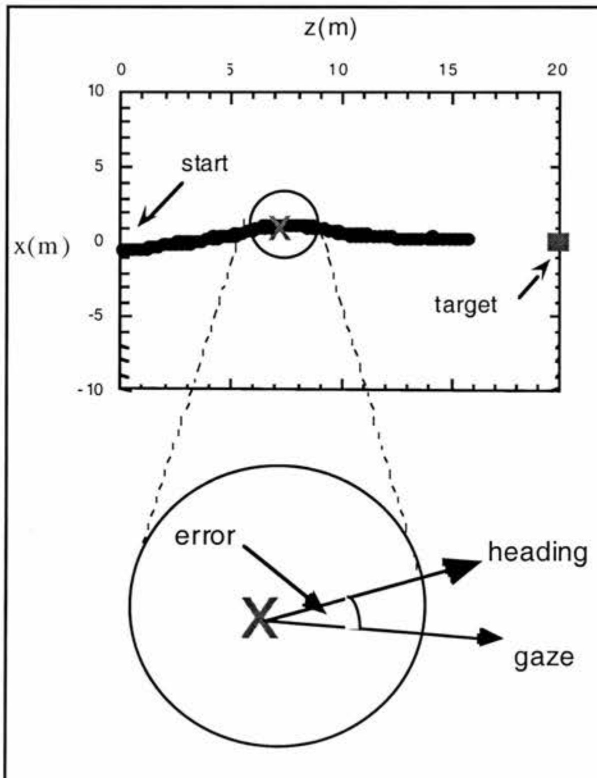


Figure 4. An example trial showing observer's progress towards the tower (top panel). The circle shows an area that has been enlarged below to indicate the angles we refer to. Heading error was defined as the difference between gaze (direction of target) and heading direction (tangent to path at each instant).

There were 25 trials per condition and the conditions were randomly interleaved. Statistical analysis and summary graphs use mean heading error over the full time course.

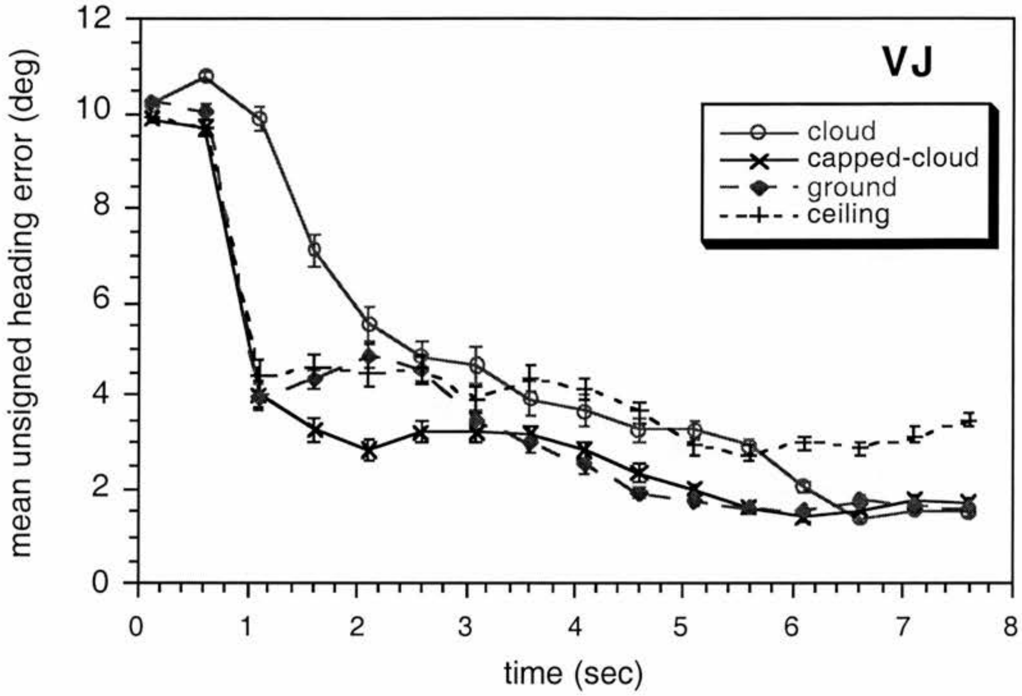


Figure 5. An example plot of results for observer VJ. Mean unsigned heading error (degrees) sampled with 200msec windows at 0.5 secs intervals across the time-course, standard errors shown. Note how error is systematically reduced from initial 10deg. Mean errors across whole trial for the above observer are: 4.632^{o} (cloud), 3.250^{o} (capped cloud), 3.507^{o} (ground), 4.377^{o} (ceiling)

Care should be taken attempting to directly compare these results and those from previous judgement tasks, see example time-course plot in fig 5.

4.2.5. Results and Discussion

4.2.5.1 Experiment 1

Our measure of performance for this task was average heading error over the trial. Heading error was defined as the difference between gaze (direction of target) and heading direction (tangent to path at each instance) (see figs 4 & 5).

An analysis of variance showed a significant main effect of condition (p<0.0001). Pair-wise comparisons revealed the cloud condition produced significantly worse performance than the other conditions (ground plane, p<0.0001; capped-cloud, p<0.0001; ceiling, p<0.0001). There was no significant difference in performance between the ground plane and capped-cloud (F(124)=0.062, NS). The ceiling condition produced a level of performance intermediate between the cloud, and the ground (p<0.0001) and capped-cloud (p<0.0001) demonstrating an interesting lower visual field advantage (Previc, 1990; Rubin et al, 1996) (see fig 6).

These results suggest that a difference in 2D layout rather than 3D layout is responsible for the decrement in performance observed with cloud displays. This finding questions the use of pictorial depth in steering from flow.

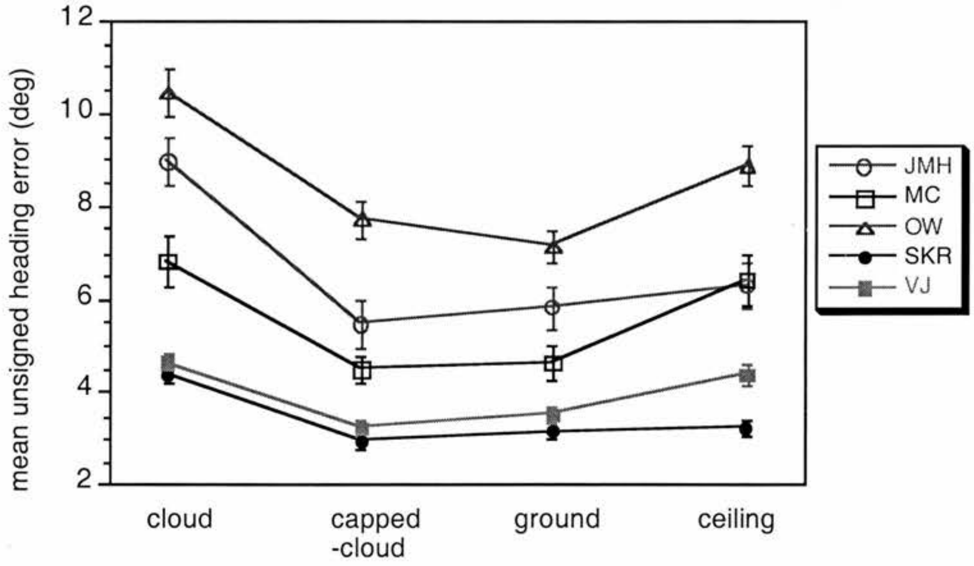


Figure 6 Interaction plot (condition & observer) of mean unsigned heading error (degrees) over time-course with standard errors indicated.

4.2.5.2 Experiment 2: Does binocular disparity aid steering?

Data were collected and analysed as for the first experiment. No difference was found between conditions (F(2,240)=0.085, NS) and pairwise comparisons revealed no simple effects (see fig 7a). Performance levels in the comparable cloud condition were remarkably similar to those in the previous 3D layout experiment. The lack of elevation of performance suggests that disparity depth is not used in steering. Failure to find a clear benefit for veridical disparity is in fact in line with that previously reported in a heading judgement task (van den Berg & Brenner, 1994b). Importantly, as performance was not impaired when depth and heading were not congruent (non-veridical stereo) we are led to conclude that the perceptual systems associated with depth and the control of heading (steering) from flow are not coupled (Shallice, 1988).

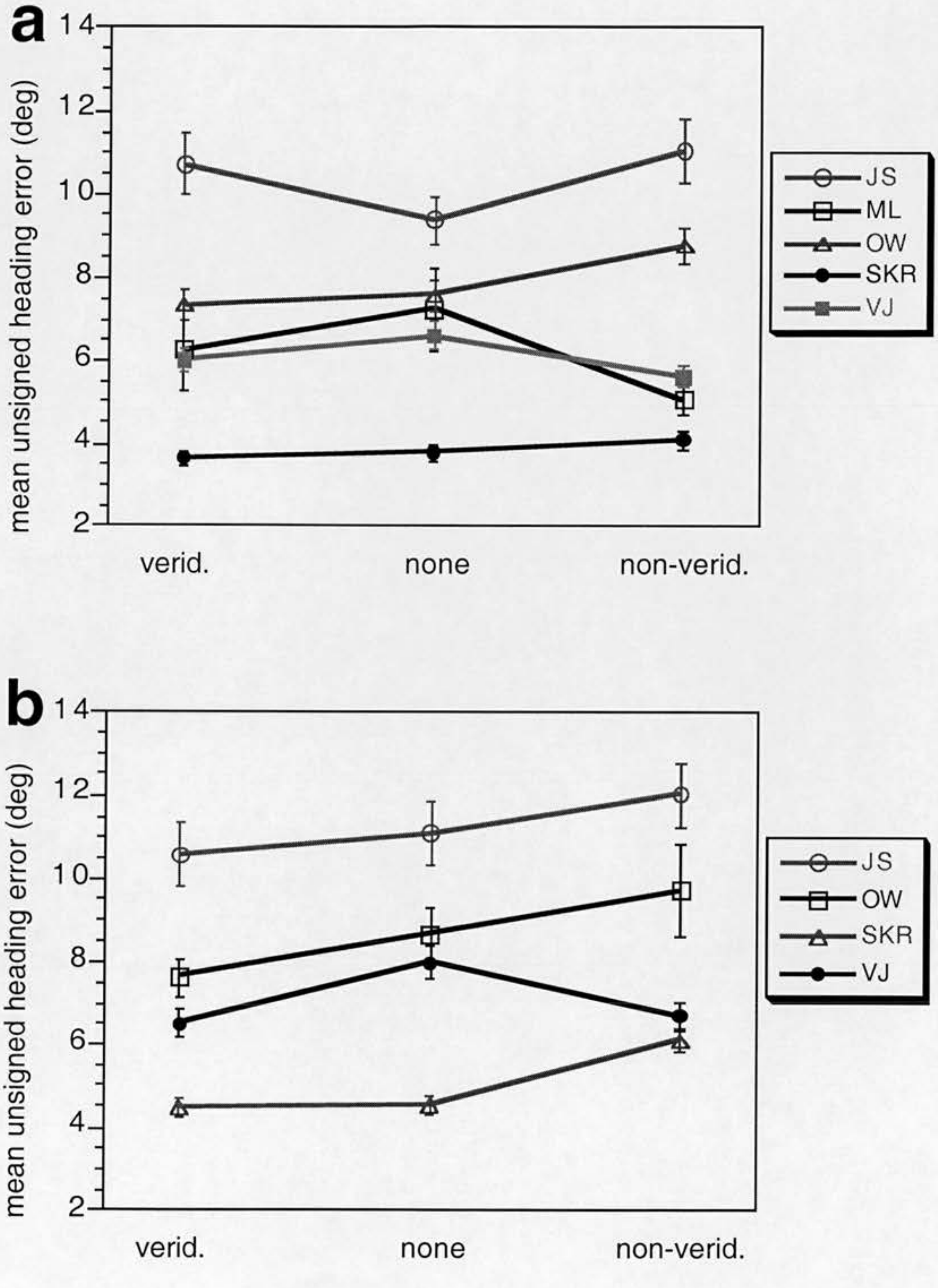


Figure 7a. No noise, interaction plot (condition & observer) of mean unsigned heading error (degrees) over time-course with standard errors indicated.
b. Noise (SNR=1), interaction plot (condition & observer) of mean unsigned heading error (degrees) over time-course with standard errors indicated

4.2.5.3 Experiment 3: Does disparity aid steering in a noisy stimulus?

This experiment was to determine if steering and heading judgements demonstrate a comparable pattern of results. Following van den Berg & Brenner (1994) we tried to open up a performance difference between the conditions through the addition of a high level of noise. When a very high level of noise (SNR=1) was added the main effect of condition almost reached significance (F(2,192)=2.965, NS) (see fig 7b). Pair-wise comparisons revealed that veridical stereo was not significantly better than the non-stereo condition (F(124)=1.953, NS), nor did the difference between non-veridical stereo and the non-stereo condition reach significance (F(124)=1.058, NS). However, a significant difference was found between veridical stereo and non-veridical stereo conditions (F(124)=5.885, $p<0.05$) (see fig 4). These results are similar to those reported previously on heading judgements (van den Berg & Brenner, 1994a).

4.2.6 Discussion

4.2.6.1 Depth gradient in capped cloud

Performance in the capped cloud was as good as for a ground plane. There are cues to depth, however, that arise for the capped cloud. The depth-range in the cloud of a point towards the bottom of the screen is more limited that the depth-range of a point towards the top of the screen (see caption for fig 1). Hence the mean, median, modal or maximum distance of a point increases from the bottom to the top of the display and this could serve as a probabilistic cue to depth. Further, points at the bottom of the screen are always near (see fig 1). In principle, an observer could deduce this and then concentrate on the movement of these points. They are potentially informative because their vectors are least contaminated by an additional rotational component due to gaze movement, however, they are very few and move very fast.

Let us reconsider the specific hypothesis that rotation subtraction uses the movement of distant objects (van den Berg, 1992). There is no part of the projected array that contains just distant objects, to assume that points near the top of the screen are distant is invalid. The top of the screen is identical to a standard cloud. Therefore we can take the results as being in conflict with the van den Berg hypothesis.

An additional point to consider is the upper/lower visual field effect (ceiling vs. ground plane). A ceiling and a ground plane contain identical depth order cues and are also both surfaces. The relative performance differences support the theory that retinal locus is important in perception and control of heading.

4.2.6.2 Rotation rates

A large number of recent studies have been concerned with rotation rates. Could it be that our result does not hold at a 'higher rotation rate' (i.e. by increasing the foot speed or initial heading direction)? This issue illustrates the difficulty in generalising from psychophysical judgement tasks to more naturalistic steering tasks: Firstly, in a steering context the rotation rates that observers are exposed to depend upon their responses. If they do not attempt to correct their heading direction then the rotation rate will increase. In an active control context, however, high rotation rates may actually be considered as **information**. From the individual results (see appendix) it can be seen that two observers delay their steering response (note the increase in error does not indicate a steering action in the wrong direction, rather it indicates a lack of steering) in the cloud condition. There may be two explanations for this, the first is that the cloud condition is difficult and so it takes longer for a percept of heading to develop or for the angle between heading and gaze to pass threshold. Alternatively, observers may be able to exploit the higher rotation rates (note that mean or total global rotation is not informative as this varies with the distance to the target and total extent of the environment).

Second, we question whether the high rotation rates that are introduced in psychophysical judgement tasks are valid for natural contexts and actions. Gaze rotation rate is a function of the instantaneous angle (α) of the fixation point from the locomotor trajectory, the distance (Z) of the observer from where a linear trajectory will pass nearest to the fixation point, and the locomotor speed V. For locomotion at a constant velocity we can consider the instantaneous time before the observer passes the fixation point as TTP = Z/V. The gaze rotation rate ω is then specified as:

$$\omega = 0.5 \sin(2\alpha)V/Z = \qquad 0.5 \sin(2\alpha)/TTP \qquad (1)$$

This illustrates that to achieve high rotation rates the fixation point must either be eccentric or the time to passage relatively short (in other words, the distance in depth between observer and fixated point should be short). Hence, a rotation rate of 5deg/s (0.0873 rad/s) will be achieved for a fixation point offset from the instantaneous locomotor trajectory by α = 5deg, but only 1sec before it is passed. For a fixation point that has moved α = 20deg from the locomotor trajectory, the rotation rate will rise to 5deg/s at a TTP = 3.68s, but within 1 sec the trajectory will have taken the observer 28 degrees eccentric to the target, and within 2 seconds the fixation point will be at α = 44 degrees eccentricity. These eccentricities are at the limits of conventional displays and probably beyond normal practice. If a car driver maintained fixation on a pedestrian for 2 seconds and until his/her gaze was at 45 degrees to the road, it would not be surprising to note that their heading perception was poor. In natural settings such rotation rates are only encountered transiently and the human gaze system is well suited to using alternating fixations to avoid the kind of sustained rotation rates used in previous psychophysical settings.

4.2.6.3 Contradictions?

Stone & Perrone (1997) questioned the data supporting the role of depth in perception of heading. They additionally presented data that showed accurate heading perception in the absence of depth even with high rotation rates. Ehrlich et al. (1998) found no significant improvement in performance at heading judgements associated with the presence of depth information. Cutting's model assumes that depth information is used in heading (indeed, differential motion parallax is only useful when the depth of elements is known). Vishton, Nijhawan & Cutting (1994) did explicitly examine the importance of depth in a study that compared performance when optical element size was consistent or inconsistent with fronto-parallel motion. They found that performance was only at chance when optical size was inconsistent. Unfortunately, the study is only documented in a the short published conference abstract and as investigators have not previously or since used similar stimuli it is not possible to guess the many missing details. Therefore, it is not possible to critically appraise that study here.

The data presented here demonstrate the same pattern as those of van den Berg & Brenner (1994b) for perceptual judgement. However, we argue that under natural conditions when steering around an environment, disparity depth has no influence on performance. With a very high level of noise, a performance difference was found but such circumstances are not regularly encountered in natural environments. Why is there an effect of depth with a high level of noise? Given the commonality of information or anatomical inputs and outputs it is unlikely that the systems for depth and heading are absolutely isolated. For instance it is likely that they both are implicated in the perception of environmental layout. If so, contradictory outputs of the two systems will increase a general level of uncertainty, consistent ones will reduce it.

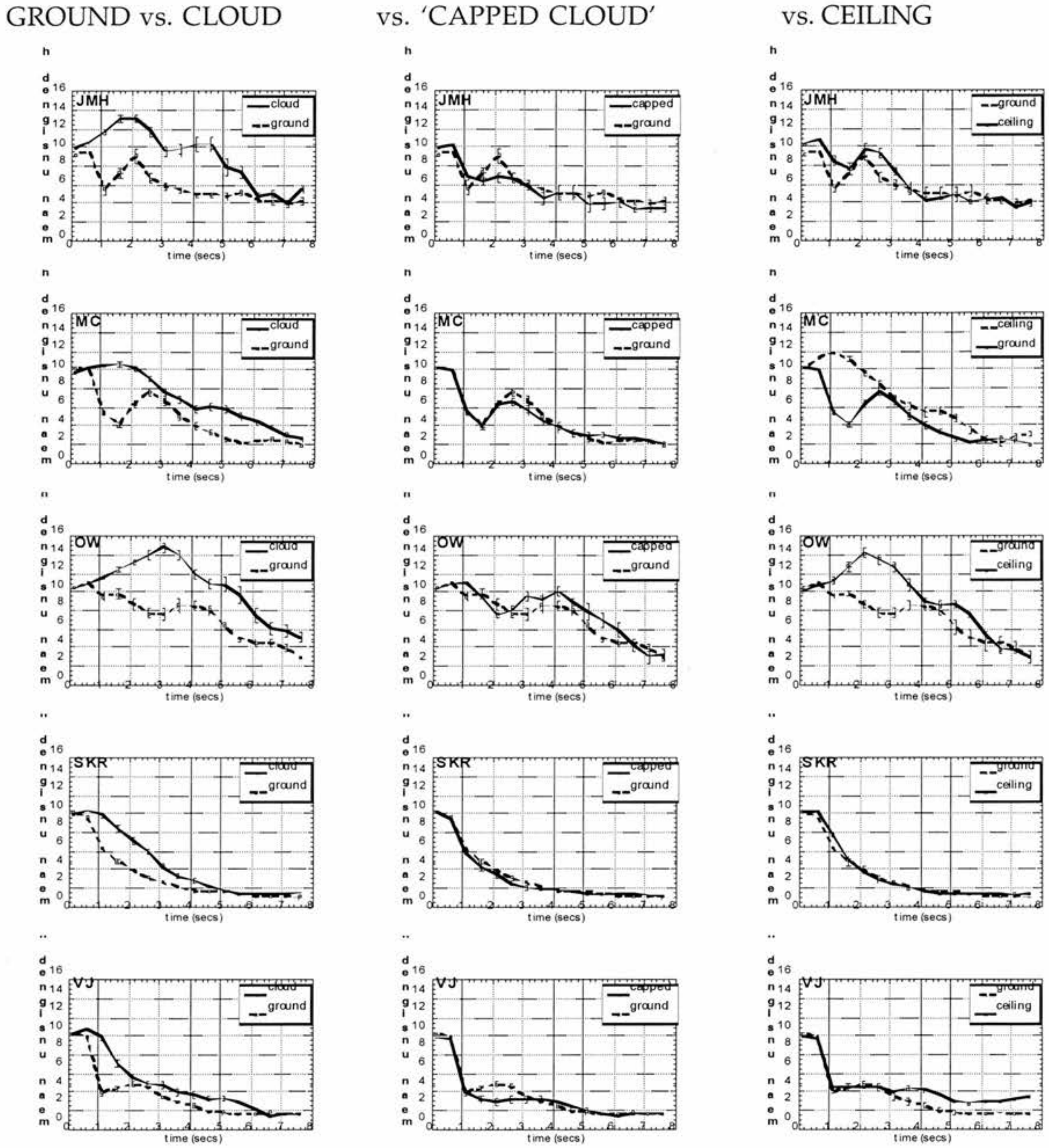
4.2.7 Summary

To summarise, although depth may be used in other aspects of locomotor control, here we find a functional independence of the perceptual systems for depth and the control of heading from flow. The presence of depth information (either height-in-scene or disparity) does not confer advantage for the perceptual control of heading (steering). Whatever strategy is used to steer using optic flow, it is not aided by providing either height-in-scene or disparity depth. Retinal motion distribution, not depth, is the major determinant of performance for the perceptual control of heading from flow.

4.2.10 Appendix

Performance curves shown over are from Rushton & Harris (1998). Graphs show individual results. Mean unsigned heading error (degrees) sampled with 200msec windows at 0.5 secs intervals across the time-course. Standard errors are shown.

4.2.10.1 Pictorial Depth and Performance


GROUND vs. CLOUD

vs. 'CAPPED CLOUD'

vs. CEILING

Note: clear divergence in performance between conditions. (Cloud upper lines - worse perf.)

Note: curves/conditions practically inseparable.

Note: some observers show a difference in performance between the ground and ceiling(worse).

Figure A1: Performance with ground plane, capped cloud, cloud and ceiling

4.2.10.2 Binocular Disparity and Performance

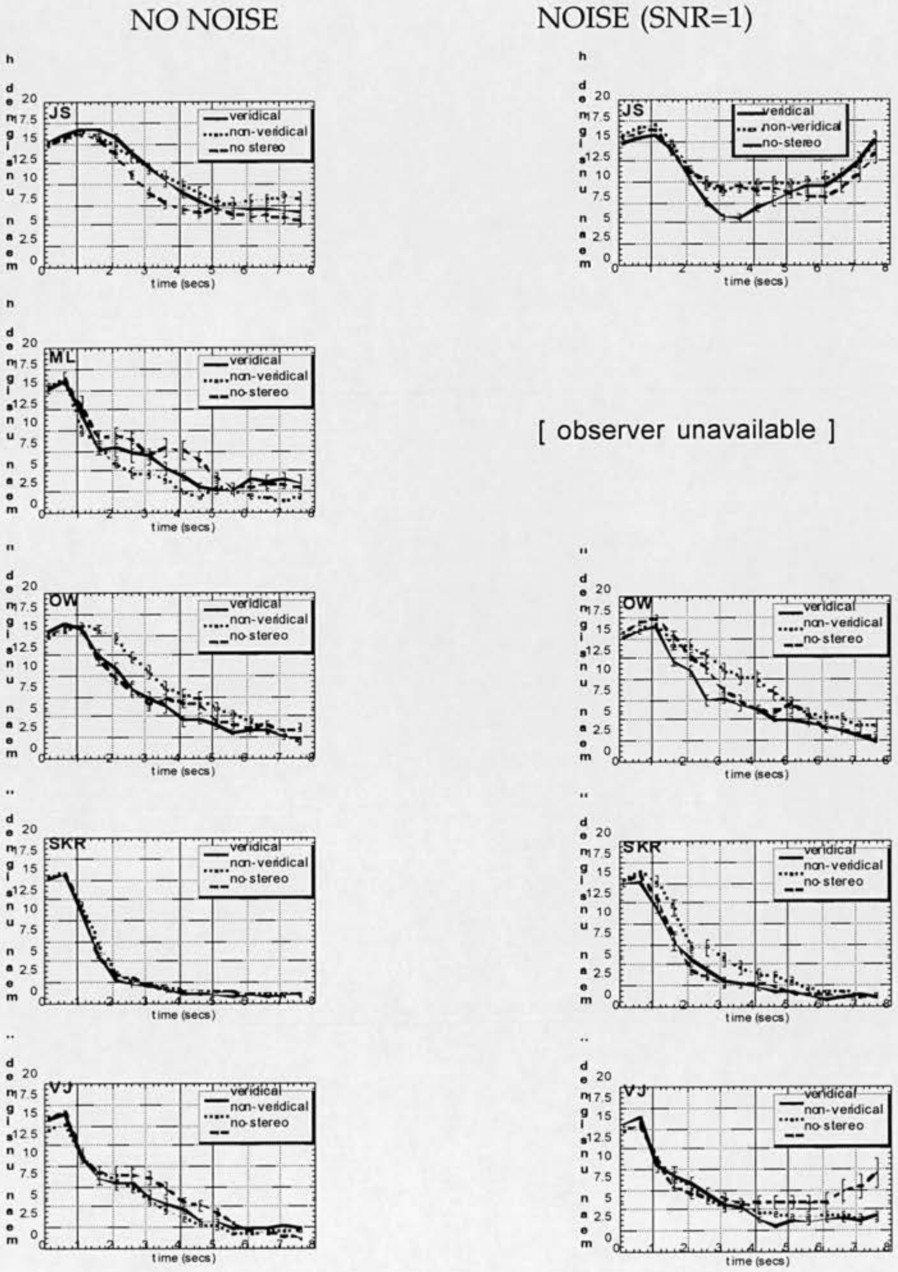


Figure A2: Performance with veridical stereo, non-veridical stereo and no stereo.

4.3 Afterword

The study was concerned with the perceptual control of locomotion rather than the vection-related percept of locomotor direction or "heading". There may be important differences between the two. Indeed a clear dissociation between the influence of depth on perceptual control and heading percept would have been suggestive of the two relying on different mechanisms[3].

Unfortunately it is difficult to conclude that a comparison of van den Berg & Brenner's judgement results against the results reported in chapter 4 indicates any such difference. First, the disparity depth results were actually rather similar (no difference unless a very high level of noise was added). My conclusion was based upon a different interpretation rather than qualitatively or quantitatively different results. Second, the pictorial depth vs. retinal motion distribution finding is likely to be replicatable with a judgement task so I do not feel confident citing that result as informative.

Therefore, the depth and perceptual control of locomotion finding cannot be taken as evidence for different mechanisms.

Some of the theoretical consequences of the findings have been discussed. The finding that depth has no role in active control of locomotor direction is compatible with the hypothesis about the use of a minimal set of information.

---

[3] Such a dissociation may have been noted in Wann, Rushton & Lee, 1995. In this study it was reported that during active control, the instantaneous velocity flow field is not the optimal input. This result contrasts with that reported by Warren et al (1992) who used perceptual judgements and similiar flow-field manipulations.

Are there any practical consequences of this finding? The original paper published on depth and perception of heading (van den Berg & Brenner, 1994) made reference to car driving. It is hard to think of any real recommendations regarding driving that can be derived from the results on steering and depth reported in this chapter. However, there are possibly consequences for the design of 'Virtual Environments' or 'Tele-operation' systems - systems that allow some to remotely guide and operate a machine (such as in the case of bomb disposal).

4.3.1 Virtual Environments and the provision of depth (from Rushton & Harris, 1998)

In the natural world humans effortlessly drive cars along twisting roads, or walk pavements crowded with pedestrians, seldom suffering a collision. In contrast, in Virtual Environments (VEs) users often clumsily crash around their environments, struggling to reach their destination. The transformation of the agile to the clumsy suggests that there is a mismatch between the user and the prototypical VE's visual content or control devices. Thus it is necessary to evaluate the contribution of different sources of information in guiding locomotion and ensure that the important ones are included when designing a VE.

There has been an interest in providing stereoscopic depth and many claims have been made for the benefits associated with it. The provision of stereo depth information in a VE system brings costs in both resources and potentially comfort. For a Virtual Environment, providing two disparate images for the left and right eyes requires up to double the computational power or a decrement in temporal or spatial resolution of the rendered images. In a tele-operation system, stereo depth requires two cameras and may introduce problems related to the necessity of rotating (or verging) the stereo cameras so as to bring the target object into alignment (Wann, Rushton & Mon-Williams, 1995).

For a HMD, stereo requires that either the left and right image pair be rendered on two separate displays or time multiplexed. This respectively removes the potential choice of using some HMDs or leads to a drop in temporal resolution. For a desktop based system the user must use shutter glasses, anaglyphs or a lenticular display with similar contrast or colour gamut costs. Additionally it has been suggested that stereo HMDs may cause greater user discomfort than bi-ocular displays (same image to each eye; Mon-Williams, Wann & Rushton, 1993; Rushton, Mon-Williams & Wann, 1994). Consequently, it is critical to consider whether stereo depth may aid self-motion.

The data reported in this chapter suggests that the provision of depth information (either 3D layout or disparity) does not confer any advantage in the perceptual control of heading. 2D motion distribution, not 3D layout, is the major determinant of performance for the perceptual control of heading.

Therefore it appears that when designing a system for **locomotion** that computational resources are better spent on improving the temporal and spatial resolution of a single display than in providing stereoscopic depth information. Providing rich flow in the lower visual field would aid most in supporting a user in steering around an environment. In virtual environments this may often be most easily provided with a textured ground plane. With a tele-operation system this may simply be achieved by titling a camera down slightly so as to bring the surface being traversed into view.

4.4 References

van den Berg, A.V. (1992). Robustness of perception of heading from optic flow. *Vision Research,*, **32**, 1285-1296.

van den Berg, A.V. & Brenner, E. (1994a). Humans combine the optic flow with static depth cues for robust perception of heading. *Vision Research*, **34**, 2153-2167.

van den Berg A.V. & Brenner, E. (1994b). Why 2 eyes are better than one for judgements of heading. *Nature*, **371**, 700-702.

Cormack, R. & Fox, R., (1985). The computation of retinal disparity. *Perception and Psychophysics*, **37**, 176 - 178.

Cutting, J.E. (1996). Wayfinding from multiple sources of local information in retinal flow *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1299-1313.

Cutting, J.E., Springer, K., Baren, P.A. & Johnson, S.H. (1992). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: Human Perception & Performance*, **121**, 41-72.

Ehrlich, S. M., Beck, D. M., Crowell, J.A., Freeman, T.C.A., & Banks, M.S. (1998). Depth information and perceived self-motion during simulated gaze rotations.*Vision Research*, **38**, 3129-3145.

Gibson, J.J. (1950). *The perception of the visual world.* Greenwood Press, Westport CT.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Lawrence Erlbaum Associates, Hillsdale, NJ.

Georgeson, M.A., Freeman, T.C.A. & Scott-Samuel, N.E., (1996). Sub-pixel accuracy - psychophysical validation of an algorithm for fine positioning and movement of dots on visual-displays. *Vision Research*, **36**, 605-612.

Goodale, M.A. & Milner, A.D., (1992). Separate visual pathways for perception and action. *Trends Neurosci.*, **15**, 20-25.

He, Z.J.J. & Nakayama, K. (1994). Perceived surface shape not features determines correspondence strength in apparent motion. *Nature*, **367**,173-175.

Land, M.F. & Lee, D.N. (1994). Where we look when we steer. *Nature*, **369**, 742-744.

Milner, A.D. & Goodale, M.A., (1996). *The visual brain in action.* Oxford University Press.

Mon-Williams M., Wann, J.P. & Rushton, S.K. (1993). Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display, *Ophthalmic & Physiological Optics,* 13, 387-391.

Nakayama, K., (1994)J. ames J. Gibson - an appreciation. *Psychological Review*, **101**, 329 - 335.

Palmisano, S. (1996). Perceiving self-motion in-depth - the role of stereoscopic motion and changing-size cues. *Perception & Psychophysics*, **58**, 1168-1176

Previc, F.H. (1990). Functional specialization in the lower and upper visual-fields in humans - its ecological origins and neurophysiological implications. *Behavioural and Brain Sciences*, **13**, 519-575

Regan, D & Beverley, K.I. (1982). How do we avoid confounding thedirection we are looking and the direction we are moving? *Science*, **215**,194-196.

Roy, J.P. & Wurtz, R.H. (1990). The role of disparity-sensitive cortical-neurones in signalling the direction of self-motion. *Nature*, **348**, 160-162

Royden, C.S., Banks, M.S. & Crowell, J.A. (1992). The perception of heading during eye movements. *Nature*. **360**. 583-585.

Rubin, N., Nakayama, K. & Shapely, R. (1996). Enhanced perception of illusory contours in the lower versus upper visual hemifields *Science*, **271**, 651-653.

Rushton S.K. & Harris J.M. (1998). Stereo Depth And The Control Of Locomotive Heading. *Stereoscopic Displays And Virtual Reality Systems V*, vol.3295, ch.47, pp.247-255

Rushton S., Mon-Williams, M. and Wann, J. (1994). Binocular vision in a bi-ocular world: new generation head-mounted displays avoid causing visual deficit. *Displays* 15, 255-260

Rushton, S.K. & Riddell, P.M. (in press) Developing Visual Systems and Exposure to Virtual Reality and Stereo Displays. *Applied Ergonomics.*

Shallice, T. (1988). *From Neuropsychology to Mental Structure.* Cambridge University Press, Cambridge.

Stone, L.S. & Perrone, J.A. (1997). Human heading estimation during visually simulated curvilinear motion. *Vision Research,* **37**, 573-590.

Turrano, K. & Wang, X. (1994). Visual Discrimination between a Curved and Straight Path of Self Motion: Effects of Forward Speed. *Vision Research,* **34**, 107-114.

Wann, J.P., Rushton S. & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments, *Vision Research,* 35, 2731-2736.

Warren, W.H. (1995). Self-Motion: Visual Perception and Visual Control. IN Epstein, W. & Rogers, S. (eds). *Perception of Space and Motion* , Academic Press, San Diego, 263-312.

Warren, W.H. & Hannon, D.J. (1988). Direction of self-motion is perceived from optic flow. *Nature,* **336**, 162-163.

Chapter 5:

The influence of perceived location when walking

5.1 Foreword

W.V. was my best patient during my unilateral visual neglect research. He was generous with his time and enthusiasm, answered all my stupid questions, tolerated all my tedious testing and tried some of my more ridiculous suggestions without too obviously raising an eye-brow. After he was discharged I saw him occasionally when he came in as an out-patient.

One day his wife told me about the holiday the two of them had taken in the north of Scotland. As I recall, they had a caravan on one side of a field and had to walk across the field to reach a shop or office. Apparently W.V. walked a strange veering trajectory. I was aware of literature on neglect patients taking strange trajectories whilst walking through doorways. The explanations were couched in terms of a "magnetic attraction" to one door post (the right door post in the case of left neglect). However the studies also mentioned collisions with objects on the left side. It was not clear how such a "magnetic attraction" could explain veering behaviour when walking across an open space.

I spent a considerable amount of time trying to work out how it would be necessary to distort an optical flow field to make someone veer. I was unable to find any satisfactory solutions. However, I realised that it was possible to predict a curving trajectory based upon mis-perception of ego-centric direction. Prisms change the perceived direction of objects. This paper to follow describes what happens to normals when they wear prisms.

The paper reports a simple task and a very simple putative model based upon perceived ego-centric target location to explain the results. In normal circumstances walking involves more than crossing an open space towards a target. A more fully articulated model of control of locomotion based upon perception of ego-centric direction is introduced in the chapter that follows. However, we start with the simple version.

After the paper I pick up a few thoughts on neglect and veering and also talk about why a hemianope might learn to veer.

5.2 Paper

from Rushton, S.K., Harris J.M., Lloyd M.L. & Wann J.P. (1998). The role of perceived location during locomotion on foot. *Current Biology,*

What visual information do we use to guide our movement through an environment? During self-movement, the direction of movement ('locomotor direction'), is specified by a point in the optical flow field from which all motion vectors radiate, (the 'focus of expansion', FoE; Gibson, 1950; Mollon, 1997; Calvert, 1950). However, if a point off to one side is fixated, then the FoE no longer specifies locomotor direction(Regan & Beverley, 1982). Models have been proposed that remove confounding rotational motion components due to eye-movements by decomposing the retinal flow into its separable translational and rotational components (see Longuet-Higgins & Prazdny, 1980 and Rieger & Lawton, 1985 as early examples). An alternative theory is based upon the use of invariants in the retinal flow field (Cutting et al, 1991). The assumption that underpins all these models (see also Hildreth, 1992 and Perrone & Stone, 1994), associated psychophysical studies (e.g. Royden, 1997 and van den Berg, 1993) and neurophysiological research (e.g. Saito at al, 1986; Duffy & Wurtz, 1991 and Britten & van Wezel, 1998) is that locomotive heading is guided by optic flow. In this paper we wish to challenge this assumption for the control of direction of locomotion on foot. Here we explore the role of perceived location. We used displacing prism glasses and recorded walking trajectories. The results suggest that perceived location, rather than optic or retinal flow, is the predominant cue that guides locomotion on foot.

We start with an observation. W.V. has unilateral visual neglect (UVN - see Shillcock et al, in-press for details of W.V.). W.V.'s wife reports that he consistently walks a peculiar veering course to objects of interest. Current theories of perception of locomotor direction based on optic flow appear unable to explain or predict W.V.'s trajectory. However, UVN is associated with the mis-perception of location. So, perplexed with the report of W.V.'s behaviour, we attempted to manipulate perceived target location for normal individuals, to see if similar veering trajectories could be induced.

Flow based theories of heading are concerned with the perception of locomotor direction relative to objects or elements in the environment or image. In the simplest case, for an eye fixed in its socket, the locomotor direction is specified by the position of the FoE within the image, for example, 5deg to the left of a target tree. When a horizontal wedge prism is placed before the fixed eye, the entire image of the world is shifted on the retina (see fig 1).

Because the **whole** image is deflected by a prism, the position of the FoE **relative** to the target tree and all other objects within the image or environment is unchanged. The perception of the locomotor direction should remain veridical (still 5deg to the left of the tree) if perception of locomotor direction relies upon optic flow. (This is also the case when locomotor direction is recovered from a more complicated flow field including an eye-movement.) For example, a simple and representative flow-field based strategy for reaching a target, the *FoE-target* strategy can be described as follows: (i) walk forward; (ii) locate target within image; (iii) locate FoE from flow within image; (iv) if the two are not coincident, then modify locomotor direction and reiterate loop. With or without prisms the *FoE-target* strategy (and all other flow-based strategies - see fig 1 caption) should lead to a straight course to a target.

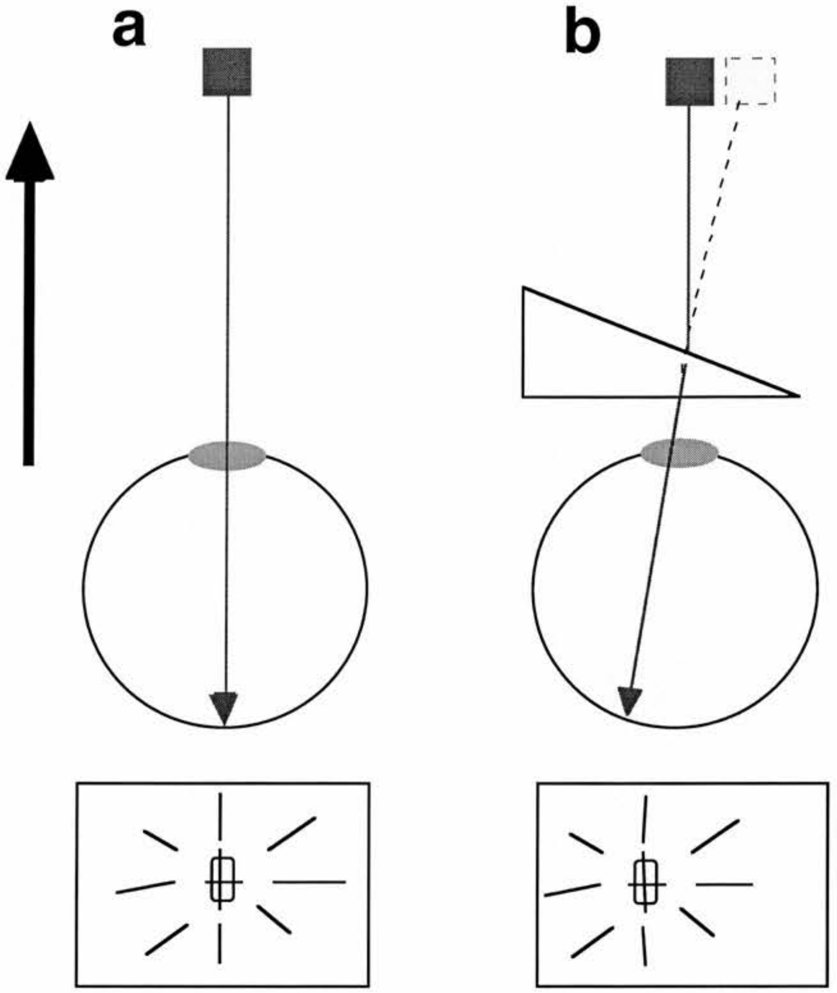


Figure 1: Retinal position of object located straight ahead of the participant: **a** without prism, **b** with prism. **lower:** instantaneous optic flow field corresponding to translation directly towards target (direction indicated by black arrow), **a** without prism, **b** with prism. Note in **a** and **b** the focus of expansion is coincident with the target, but in **b** both the target and FoE are displaced to one side. Therefore, if the **relative** position of the FoE and the target is used to control locomotion on foot then the participant should walk directly towards the target. Note, changing to ego-centric co-ordinates and hypothesising that independent neural systems are responsible for determining the position of the FoE and the target does not change the prediction. Such a model would predict a prism-induced error in both the co-ordinates of the FoE and the co-ordinates of the target. This constant error would cancel out leaving the correct **relative** position of the FoE with respect to the target.

However, a prism changes the perceived location of an object relative to the mid-line of the body (the locomotor axis) (Rock, 1966). If location guides locomotion on foot, then placing prisms in front of the eyes should perturb perception and control of locomotor direction. For example, a simple strategy based upon perceived location is the *perceived-direction* strategy which can be described as follows: (i) walk forward; (ii) rotate gaze to fixate the target; (iii) rotate body in direction that should reduce angle between gaze and mid-line; (iv) evaluate difference between angle of gaze and orientation of the body and reiterate loop. Under normal circumstances the *perceived-direction* strategy will work successfully and result in a straight course to the target. However, if the person consistently misperceives the location of an object relative to their body (as happens when wearing prisms) or misperceives the mid-line of their body (as often happens after brain injury) then they will mis-align their locomotor axis with the true direction of the target and so produce a constant heading error. Thus, a person wearing prisms and using this strategy should walk a veering trajectory (see fig. 2)

Perceived Direction and Reference Frames

The direction of an object can be described in a variety of coordinate frames. Environmental or spatiotopic frames describe the direction of one object with respect to an environmentally defined axis.

When considering human or animal interactions it is more useful to consider the direction in a body (or ego) referenced coordinate frame. There are a number of possible frames: 'retinotopic' (directions relative to the fovea), 'head-centric' (directions relative to the head), 'body-centric' (directions relative to the trunk). Head-centric directions require the combination of retinal coordinates and direction of the eye in the head. Body-centric directions require the combination of retinal coordinates and gaze direction (head on should + eye in head).

Under some circumstances perceived direction and veridical direction may differ. For instance the direction of an eccentric object may be mis-judged due to a perceptual compression or dilation of space. Alternatively, manipulations may perturb it: e.g., perceived body-centric direction may be changed by injecting cold or warm water into an ear, wearing prisms, following perceptual adaptation or following acquired brain injury.

Location is given by direction and distance (in a polar coordinate system) and when the term location is used it is intended to indicate that distance as well as direction may be of importance.

The perceived-direction strategy suggests that the observer attempts to place a target object directly 'in-front' of themselves. The relevant reference frame for locomotion is the body-centric frame (the orientation of the trunk with respect to the environment normally provides the best indication of the direction of travel). Therefore, an observer attempting to keep a target object 'straight-ahead' or 'in-front' will try to place the target object on the axis of the body mid-line.

see sections 6.2.3 through section 6.2.5 and the associated figures.

We set out to examine the respective influence of the flow-field and perceived location in guiding locomotion. Participants wore glasses with wedge prisms or Fresnel prisms deflecting right or left. An experimenter held out a target ball and asked the participants to walk over and touch the ball. The participants walked at a brisk pace for approximately 10 to 15 m. The participants' trajectories were recorded by a camera 33 metres overhead. Video frames were captured on a PowerMac 8500 computer and digitised using the public domain NIH Image program.

The trajectories taken by the participants follow a curved path similar to the *perceived-direction* prediction. From the digitised data it is possible to determine the locomotor direction at any point during the person's motion (tangent to the curve of their path), the direction of the target, and the difference between them, which we define as the target-locomotor direction error (α). The simple locomotor direction strategy based upon perceived direction predicts that α should be equal to the angular deflection of the prism.

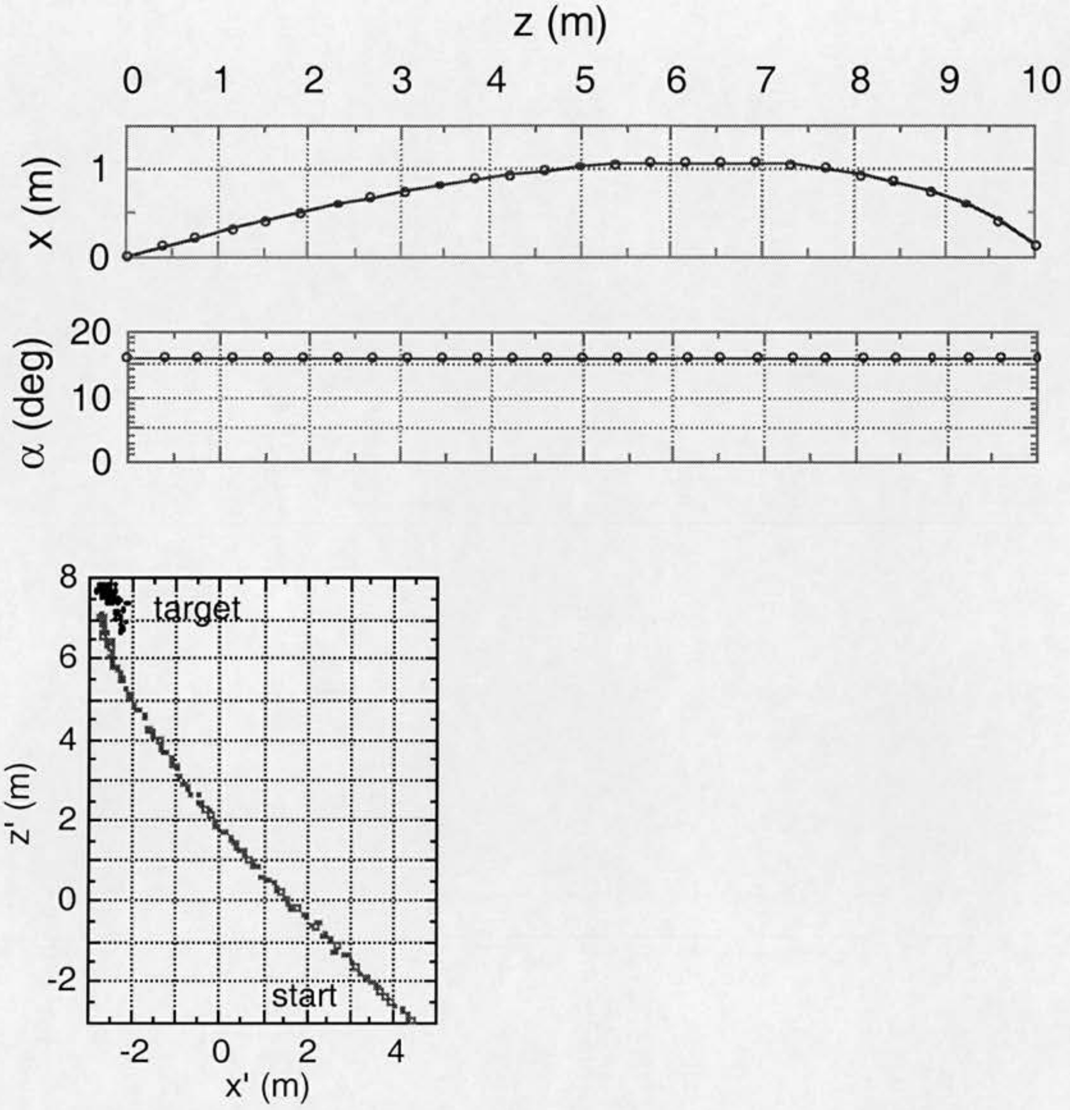


Figure 2 **Upper panels.** Simulation of the trajectory and direction error when wearing prisms by a simple model using target direction, rather than optic flow. x and z are distances parallel and perpendicular, respectively, to the starting position of the participant (facing along the z-axis). The **top panel** shows a plan view of the predicted trajectory of a prism-wearing participant walking in the perceived direction of the target (perceived direction is offset from actual position by the angular deflection of the prism glasses). The simulation plot is for a 16deg deflection which is the approximate angular deflection of the wedge prisms used in our experiments. The **middle panel** shows the angle, α, between the instantaneous direction of the target and the direction of locomotion (tangent to the curve) which remains constant throughout the trajectory. The trajectories taken by the participants were similar to that shown in the upper panel. An independent replication of the prism manipulation also produced curved trajectories (Brian Rogers, *personal communication*). Additionally, it was found that participants could partially straighten their path by explicitly trying to use motion parallax. However, if this strategy was used it was noted that "the feet keep trying to do something different". If the participant adapted to the prisms then upon their removal they would veer in the opposite direction. These observations concur with our own, it feels strange and unnatural to use such a deliberate motion-parallax strategy and a participant trying to do so can easily be identified by their peculiar gait and body twisted at the waist. **Bottom panel:** A representative trajectory (participant 3, wedge prism, going right in Figure 3) of an observer (shown in red) approaching a target (blue). The plot shows raw digitised data, with axes x' and z' showing distances in camera co-ordinates.

Fig. 3 shows values of α as the trial progresses. In general, α is close to the prism deflection angle (as predicted by the *perceived-direction* model) and clearly not close to zero as a flow-based model would predict.

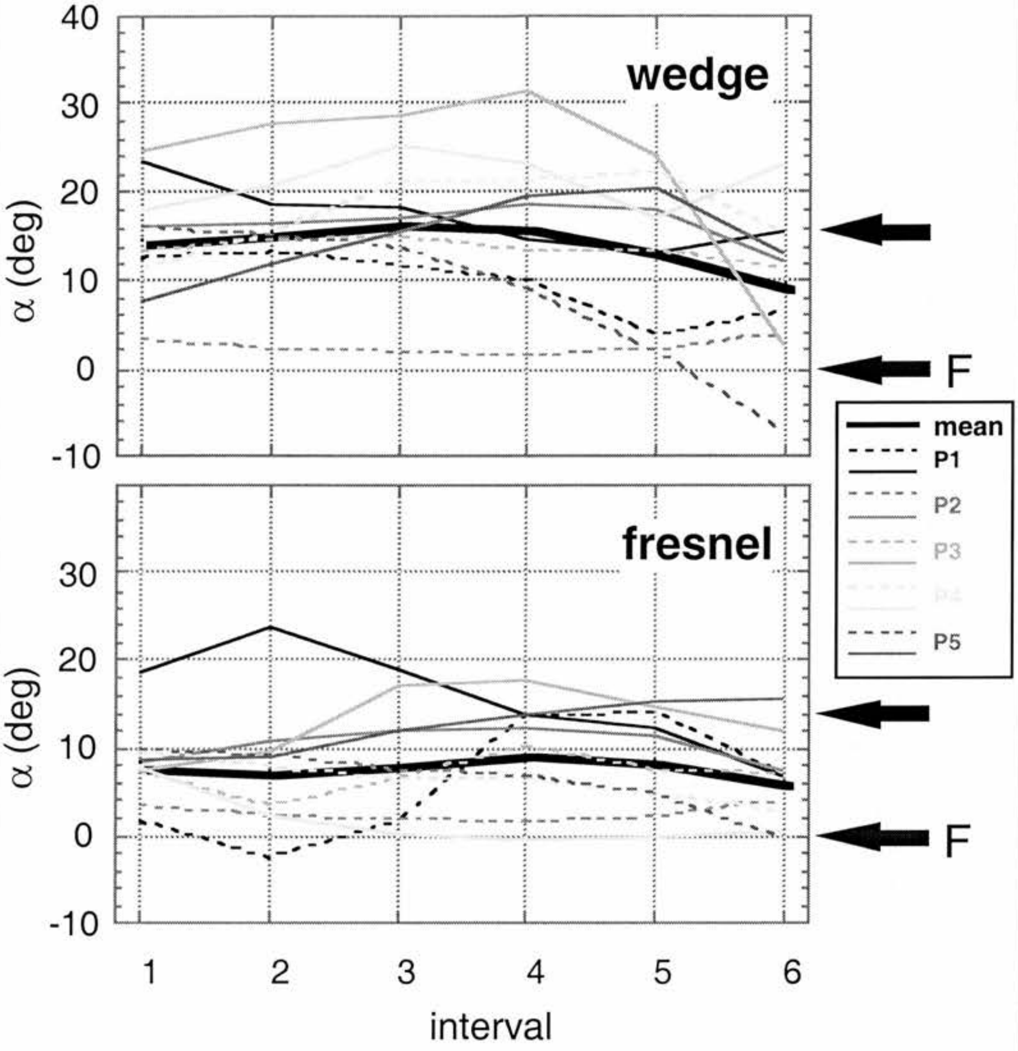


Figure 3. Mean target-heading angle (α) across trial for 5 participants. Raw trajectory data was smoothed with a gaussian, σ = 8 data points (1/3 sec), window = 16 data points, (2/3 secs). Trials were normalised and divided into 6 intervals (by distance; mean = 12.7m [1.3]). The upper panel shows α when wearing wedge prisms (~16 deg angular deflection), lower panel when wearing Fresnel prisms (~14 deg angular deflection; wider field of view, lower optical quality). Full lines = left deflection, dotted lines = right deflection. The thick black line on each plot shows the mean across all participants and both deflection directions. Black arrows show the prism deflection (*perceived-direction* prediction) and blue arrows the *FoE-target* prediction.

In general, the error, α, was fairly constant across the whole trial for both types of prism. For the wedge prisms, α was close to the value of the prism deflection (black arrow). The Fresnel prisms produce proportionately less veering than wedge prisms. The Fresnel prisms have a wider field of view and we hypothesise that this may be important - not because more peripheral flow can be seen but rather that parts of the body, in particular the nose, are visible (Gibson , 1979) and this may serve to attenuate the effect of the prisms on mis-perception of ego-centric directions.

It could be hypothesised that flow is used in our task, but that some time is required before it can be used. If so, participants might have started on the wrong trajectory and then needed several seconds to perceive direction of locomotion from the flow and act on it. Such a situation would also predict a non-straight trajectory. A late correction to the trajectory would show up in our analysis as a dramatic reduction of α at some time during the time-course. Specifically, α should reduce to zero after a second or so (about $1/7^{th}$ of the way along the trajectory). This is clearly not the case as even the longest estimate of a locomotor reaction-response time (see Cutting et al, 1991) would predict that locomotor direction would be corrected before half the trial distance is walked. Nonetheless we thought it informative to see if we could demonstrate that locomotor direction was continuously, rather than periodically, regulated.

In a second set of trials we moved the target into or away from the path of the participant whilst they were walking. If locomotor direction is continuously controlled, for the *perceived-direction* model the prediction is the same as that for a stationary target, namely that α retains an approximately constant value throughout the trial. The path will be clearly dissimilar from a predicted *FoE-target* trajectory.

Calculation of α shows that it remains approximately constant throughout the trial (mean of first interval = 11.6deg, mean of last interval = 11.2deg). This is similar to the results when using a stationary target and compatible with the use of a perceived direction strategy and continuous regulation of direction of locomotion.

The results were similar to those for the previous static target set. Target-heading angle, α was found to remain approximately constant throughout the trial (fig 4 shows a plan view of two trials), as predicted by the *perceived-direction* model. Thus, direction of locomotion is controlled on-line, in a continuous manner.

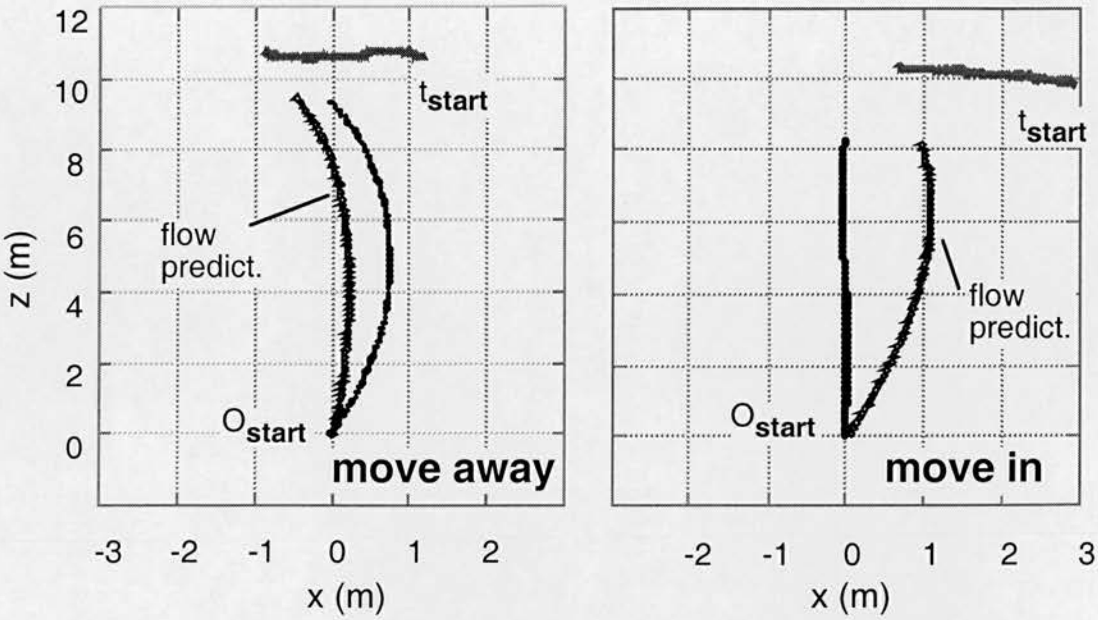


Figure 4. Two representative trajectories of a prism-wearing participant approaching a moving target. The red line indicates the trajectory of the target, the black line the trajectory of the participant. The blue line indicates the predicted trajectory that would result from using flow and walking in the instantaneous direction of the target. It is clear that participants do not follow trajectories that would be predicted from using flow. **Left panel:** participant initially walks to the right of the target and the target starts to move left, away from the participant's path. **Right panel:** participant initially walks to the left of the target and the target moves left, into the participant's path.

Our evidence that locomotor direction is regulated continuously further undermines a possible flow latency hypothesis, as was also rejected by our consideration of the static target data presented above. It is parsimonious to conclude that when moving on foot, a person's trajectory is predominantly controlled by the perceived location of a target relative to the body. This is an efficient and economical solution as knowledge of the orientation of the body with respect to objects is necessary during interception or passing.

We should note that many studies have required judgements of locomotor direction after a second or so of viewing a simulated translation through a projected abstract environment. These previous experiments have shown that people can determine their direction of locomotion from a flow field. Therefore, as people possess the ability to determine direction of locomotion from flow it appears likely that they will exploit it in some situations. However, our study reveals that a person walking through a real environment appears to be primarily influenced by perceived location (see Llewellyn, 1971 for a similar conclusion based upon target drift).

To return to W.V., do these results help us account for his veering walks? We believe they may. It has been reported that some patients with UVN mis-perceive their mid-line (Karnath, 1994). If W.V. perceives his mid-line as being away from its true position, then when he places the target apparently 'straight-ahead' and walks forward, he will not be walking towards the target. Therefore, we would predict a similar shape of trajectory to that shown by normals wearing prism glasses. The possibility also arises that the veering of an individual with U.V.N. may be nulled through the use of prism glasses. However, we must wait for another suitable patient to test this model of UVN veering as fortunately, W.V., has learnt to walk in a straight line. Interestingly W.V. now walks in a straight line even when wearing prisms - maybe he has learnt to determine locomotor direction from the flow field.

5.3 Afterword

In this section I discuss three matters. First, I anticipate a couple of questions about the prism experiment and discuss possible alternative explanations for the results. Next, I discuss the way that the perceived-direction hypothesis for veering neglect patients could be tested. Finally I mention a further patient group, hemianopes, and explain why they may learn to walk veering trajectories.

5.3.1 Alternative explanations?

5.3.1.1 Flow equalisation?

Prisms occlude a small, highly eccentric, part of the visual field on one side and increase the field slightly on the other. In bees, locomotor direction is controlled by equalising the amount of flow in the 2 halves of the visual field (Srinivasan et al, 1991). If flow was summed over the whole of the left and the whole of the right hemi-fields, then a crude equalisation strategy could be slightly perturbed by the wearing of prisms. There is recent evidence that humans can use an equalisation cue, but the influence of the cue is readily attenuated by other flow information such as the focus of expansion (Duchon & Warren, 1998). Thus it appears very unlikely that use of such a strategy could explain our results.

5.3.1.2 Cue-conflict?

In contrast to most studies of perception of locomotor direction, participants locomoted by natural walking, through a natural environment. However, the study inevitably introduced some 'cue-conflict'. Are there grounds to suspect that participants abandoned normal control strategies and behaved differently because they noticed the effects of the prisms? We believe not. First, the 'conflict' was small. All flow-based information remained congruent and veridical, the prisms perturbed only perceived location, a previously un-recognised cue. Second, participants did not appear to have problems, they walked at a brisk pace, showed no hesitancy and little awareness of their peculiar trajectory. This behaviour can be contrasted against the 'conflict' associated with trying to consciously override the influence of prisms: in an informal replication of our study (Brian Rogers, *personal communication*), participants tried explicitly to use motion parallax (the relative motion between objects in the environment) to guide themselves. This was partially successful, but it was noted that "the feet keep trying to do something different". These observations concur with our own: it feels unnatural to use such a deliberate motion-parallax strategy and a participant trying to do so can be easily identified by their odd gait with their body twisted at the waist. In summary, problems relating to cue conflict and unnaturalness of the task or visual environment were minor in this study.

5.3.2 Neglect Veering?

The study reported above was prompted by a report that a patient with unilateral visual neglect veered. As mentioned, when I retested the patient he no longer veered. Is it possible to derive some specific, testable predictions regarding veering, mis-perception of mid-line and unilateral visual neglect? I believe so and in the following (excerpts taken from a paper awaiting patient data before resubmission, Neglect veering: An Hypothesis; Rushton) I make the case.

5.3.2.0 Introduction

It has been previously reported that some patients with unilateral visual neglect (UVN) collide with objects on their neglected side (Robertson et al, 1994). It has also been reported that some UVN patients mis-perceive the direction of their mid-line and ego-centric direction of objects (e.g. Karnath, 1994) . I suggest a simple model that may link these two findings.

5.3.2.1 Neglect Veering

Several papers mention that mobile patients with UVN veer and collide with objects on their neglected side. Why should this be so? Existing theories of locomotor heading can not explain this behaviour (see fig 5). Robertson et al (1994) suggest an attentional account: in the case of left UVN, patients are drawn to objects on the right and may collide with objects on the left because they do not notice them, or because of a tendency for patients in wheelchairs to push with their right hand and so veer left. I propose an alternative.

5.3.2.2 Ego-centric direction

It has been reported that some patients with neglect mis-perceive their mid-line, or the point directly ahead of them. For example, Karnath (1994) asked three patients with left UVN to indicate with a laser pointer, the point 'straight ahead' of their bodies. He found that his patients indicated a position approximately 15 degrees to the right of their true mid-line. Similar findings have been reported by Chokron & Imbert (1995) and Heilman, Bowers & Watson (1983) amongst others. May this mis-perception of ego-centric direction explain the veering?

5.3.2.3 Veering UVN patients?

If normals mis-perceive their mid-line or the ego-centric direction of objects veer (this chapter), might UVN patients who mis-perceive their mid-line veer for similar reasons? If a UVN patient mis-perceives their mid-line then they should also mis-orient themselves whilst walking towards objects. They too should take a characteristic veering trajectory. Figure 1 below shows the predicted trajectory.

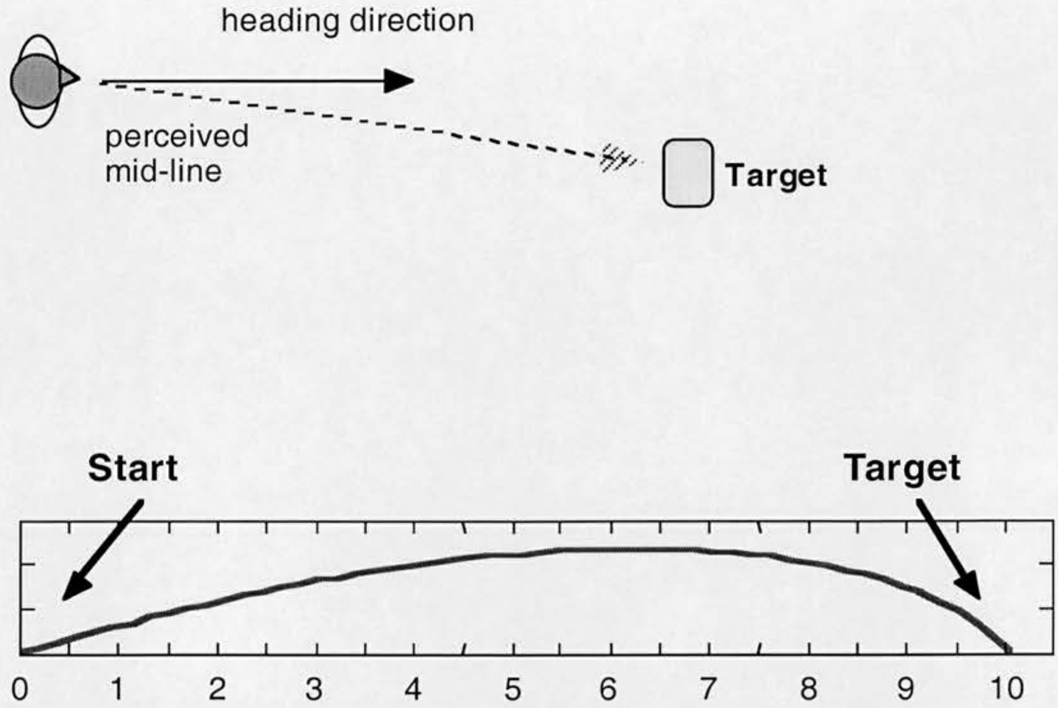


Figure 5. Top Panel: The predicted trajectory of a patient who mis-perceives their mid-line (this is the same trajectory as taken by prism wearing normals). If they walk by orienting themselves so that their target (e.g. a right doorpost) is directly 'in front' of them, then they will head off to one side. **Lower panel**: The predicted trajectory of a patient who mis-perceives mid-line. Note at any point along the course the difference between instantaneous heading (tangent to the curve) and direction of the target is equal to the error in mis-perception of mid-line. In the above example the error in mid-line perception is 15 degrees too far right. Note that it predicts patients will veer left. So if we take the case of walking through a doorway then a left UVN patient would walk towards the right doorpost but **veer out left into potential obstacles** along the way.

A further prediction that follows from the veering hypothesis is that it should be possible to stop veering UVN patients from veering by giving them prism glasses to shift their perceived mid-line back to its true position.

5.3.2.4 Conclusions

There is no currently published data to support or reject this hypothesis, however it is very simple to test. First only a subset of patients with neglect show a mis-perception of mid-line. It should be these patients that (if mobile) exhibit the veering and collisions. Second, the trajectory that is predicted is very characteristic, with (in the case of left UVN) a veer out to the left and so should be readily observable. Third, it should be possible to inhibit the veering behaviour by prescribing prism glasses.

5.3.3 Veering Hemianopes?

There is a further class of patients who may also veer that need to be examined, hemianopes. They are worthy of study for two reasons: first, to ensure that any conclusions drawn from the behaviour of patients with UVN is not confounded by effects of visual field deficits; second, because they may be interesting in their own right.

If an observer changes direction so as to maintain their target at a fixed ego-centric direction then they will reach their target. If the ego-centric direction that they hold is eccentric, e.g. 15degs, then they will walk to their target by a curving trajectory.

A hemianope has no vision in one hemi-field. It is must be unnerving to approach a target unable to see what is to one side of you. If the target is fixated during approach then this will occur and the observer may find themselves brushing against objects on their blind side. They could reduce the likelihood of this by rapidly glancing between their target and the blind side of their body. However, there is another alternative that they may choose to adopt. Through holding their target at an ego-centric direction of say 10 or 15degrees they can approach their target whilst being able to see potential obstacles to both the left and right of their path.

It would be interesting to determine if hemianopes do indeed learn to adopt a veering trajectories to reach their targets.

5.4 References

van den Berg AV. (1993). Perception of Heading. *Nature* 365: 497-498.

Britten, KH & van Wezel, RJA. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nature Neuroscience* 1: 59-63.

Calvert ES:. (1950). Visual aids for landing in bad visibility, with particular reference to the transition from instrument to visual flight. *Transactions of the Illuminating Engineering Society, London* , 15: 183-219

Chokron, S & Imbert, M. (1995). Variations of the egocentric reference among normal subjects and a patient with unilateral neglect. *Neuropsychologia,* 33, 703-711.

Cutting JE, Springer K, Baren PA & Johnson SH. (1992). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: Human Perception & Performance* 121: 41-72.

Duchon, AP & Warren, WH. (1998). Interaction of two strategies for controlling locomotion. *Investigative Ophthalmology and Visual Science* , 39, S892.

Duffy, CJ & Wurtz, RH. (1991). Sensitivity of MST neurons to optic flowstimuli. I. A contiunuum of response selectivity of large-filed stimuli. *Journal of Neurophysiology* 65: 1329-1345.

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

Gibson JJ. (1979). The ecological approach to visual perception. Lawrence Erlbaum Associates Inc, Hillsdale, New Jersey.

Heilman, K.M., Bowers, D. & Watson, R.T. (1983). Performance on hemispatial Pointing Task by Patients with Neglect Syndrome, *Neurology*, 33, 663-664.

Hildreth, EC. (1992). Recovering heading for visually-guided navigation. *Vision Research* 32: 1177-1192.

Karnath H. (1994). Subjective body orientation in neglect and the interactive contribution of neck muscle and proprioception and vestibular stimulation. *Brain* 117: 100-1012.

Llewellyn KR. (1971). Visual guidance of locomotion. *Journal of Experimental Psychology* 91:245-261.

Longuet-Higgins HC & Prazdny K. (1950). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B* 208: 385-397.

Mollon, J. (1997) "....on the basis of velocity cues alone": some perceptual themes. *Quarterly Journal of Experimental Psychology,* 50A: 859-878.

Perrone, JA & Stone, LS. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision Research* 34: 2917-2938.

Regan D & Beverley KI. (1982). How do we avoid confounding the direction we are looking and the direction we are moving. *Science* 215: 194-196.

Rieger JH & Lawton DT. (1985). Processing differential image motion. *Journal of the Optical Society of America.* A2: 354-360.

Robertson, I.H., Tegnér, R. Goodrich, S.J. & Wilson, C. (1994). Walking Trajectories and Hand Movements in Unilateral Left Neglect: A vestibular hypothesis. *Neuropsychologia,* 12, 1495-1502.

Rock I. (1966). *The Nature of Perceptual Adaptation.* Basic Books, London.

Royden, CS. (1997). Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *Journal of the Optical Society of America A.* 14: 2128-2143.

Royden CS, Banks MS & Crowell JA. (1992). The perception of heading during eye-movement. *Nature* 360: 583-585.

Rushton S.K., Harris J.M., Lloyd, M.L. & Wann, J.P. (1998). The Control Of Locomotor Heading On Foot: The Role Of Perceived Location And Optical Flow. *Investigative Ophthalmology and Visual Science,* 39, S893.

Rushton, S.K., Harris, J.M., Lloyd, M.L. & Wann, J.P (in-press). The role of perceived location during locomotion on foot. *Current Biology*

Saito H, Yukie H, Tanaka M, Hikosaka K, Fukada K Y & Iwai, E. (1986). Integration of direction signals of image motion inthe superior temporal sulcus of the macaque monkey. *Journal of Neuroscience* 6: 145-157.

Shillcock R, Kelly ML & Monaghan P. (in press). Processing of palindromes in neglect dyslexia. *Neuroreport* 9.

Srinivasan, MV, Lehrer, M, Kirchner, W & Zhang, SW. (1991). Range perception through apparent image speed in freely-flying honeybees. *Visual Neuroscience* , 6, 519-535.

Warren WH, Hannon DJ. (1988). Direction of self-motion is perceived from optical-flow. *Nature* 336: 162-163.

Chapter 6:
An eccentric model of walking

6.1 Foreword

The last chapter reports the results of a simple study and presents a very simple model. It seems likely that under more natural circumstances, the model may fall a little short of actual behaviour. In this paper I develop a more complete model of walking based upon perceived ego-centric location.

I discuss two datasets, the prism dataset introduced in the last chapter and also a dataset collected by Warren (Warren & Kay, 1997). To my knowledge this dataset is only published as an extended abstract. I have had the advantage of reading a draft paper and also talking to Bill Warren but I believe all the information on his dataset is contained in his published abstract. It may seem peculiar that I pay so much attention to his dataset. The reason for this is simple: many people know of it and the challenge I am always posed after giving a talk or submitting a paper for review is to account for his dataset. I have given up complaining that it is unpublished and have instead addressed it head on.

Interestingly, I have recently heard from J. Duysens at Nijmegen University, who has concluded that children walking on treadmills do not use optic flow. I don't currently have any more information than this.

6.2 Paper

Rushton, S.K., & Harris J.M. (in preparation). An eccentric theory of control of locomotor direction: extending Llewellyn's (1971) model.

6.2.0 Abstract

It is almost universally accepted that humans guide their movement through an environment by picking up patterns in the optical flow field. In 1971, Llewellyn asked observers to judge their simulated locomotor direction and the focus of expansion from a radial flow pattern corresponding to approach to a wall (Llewellyn, 1971). He found that observers were very poor at making such judgements. This result is clearly incompatible with the observation that humans can guide themselves very skilfully around natural environments . To account for this discrepancy, Llewellyn suggested that humans do not judge locomotor heading but rather use 'drift cancellation' (keep their target at a fixed ego-centric direction). Subsequently, it has been shown that approach to a wall is a special case, and that locomotor direction can be discerned from a flow field with some accuracy when there are depth variations in the scene (see Warren, 1995 for a comprehensive review). As a consequence, Llewellyn's theory has largely been forgotten along with his data. Our attention was drawn to Llewellyn's theory following our own recent work on perceived location and locomotion (Rushton et al, 1998a). Here, we re-examine and then expand Llewellyn's theory. Then we describe a range of approach trajectories, and the factors that may influence their use. Finally, we consider how well the extended theory and the rival flow theories account for our own recent data and seemingly contradictory findings by others. We conclude that evidence suggests that humans do not require optic flow information when walking. A model based upon perceived location can account for the experimental evidence on control of approach.

6.2.1 Overview

Gibson's work (Gibson, 1950) is usually cited as the seminal literature on the optic flow field and its role in guiding locomotion. Others such as Grindley (cited in Mollon, 1997) and Calvert (1950) appear to have competing historical claims. Whatever the parentage, the concept of the optical flow field has proved immensely compelling. A wealth of research has been conducted into identifying flow field invariants that could guide locomotion, testing the sensitivity of human and animal visual systems to flow fields, building computational models and exploring neural mechanisms. The role of the flow field in guiding locomotion has become a true 'textbook example' (e.g., Bruce, Green & Georgeson, 1996) and the topic has become one of the most high profile in vision science (see Warren, 1995 for a good review).

In this paper we show that it is possible to largely discount the role of optic flow in the guidance of locomotor direction on foot, i.e. when walking. We take a different starting point, the work of Llewellyn from the beginning of the 1970s. Llewellyn's model is briefly outlined and we describe a class of trajectories that result from its use and formalise and extend it (the 'eccentricity model'). Then we consider influences and constraints on a person's path choice and introduce the notion of perceptuo-motor 'difficulty'. Next, we examine two datasets that have been recently reported (Rushton et al, 1998a; Warren & Kay, 1997). We find the first supportive of the eccentricity model. The second has been interpreted as providing evidence against the eccentricity model. We re-examine it and suggest that it is largely compatible with the model and suggest further experimentation that should clarify matters. Finally, we consider locomotion in a naturalistic context and the constraints of obstacle avoidance, path preference and orientation.

6.2.2 Llewellyn (1971)

Llewellyn presented observers with flow patterns corresponding to approach towards a wall. Observers were asked to indicate either their direction of travel or the focus of expansion. The observers' performance was found to be very poor. Other investigators have reported a similar result (Johnston, White & Cumming, 1973; Regan & Beverley, 1982). However, approach towards a wall has been shown to be a special case (see Cutting, 1986, pp 154-161 for discussion) and under more typical experimental conditions, when depth variations in the stimulus display are present, performance is dramatically better. For example observers can judge direction of heading with an accuracy of less than a degree in the presence of a ground plane (see Cutting, *ibid.* or Warren, 1995).

At the time Llewellyn conducted his research, the importance of 3D structure was not appreciated. Llewellyn attempted to reconcile the inability to judge direction of locomotion from optic flow with observers' ability to navigate safely round the natural world. His data led him to conclude that moving observers do not use optical flow for navigation. Instead, Llewellyn proposed an alternative target-directed strategy for the perception and control of direction of locomotion. This can be summarised as follows: Whilst moving towards a target note the ego-centric direction (the angular direction of the target measured relative to your body mid-line) of the target. If during translation the target appears to drift towards the left relative to your body, then you are travelling to the right of the target (and vice-versa). If the target drifts slowly then the error in locomotor direction is small. If the target drifts fast then the error is large. To reach your target all you need to do is stop the target from drifting. This is referred to as 'drift cancellation'.

6.2.3 Trajectories: implementing a simple model

Llewellyn pointed out that if a target is held at a fixed ego-centric direction, (i.e. target drift is zero) then the observer will reach the target. It follows that if the target is held directly ahead of the observer they will reach the target by a straight path. If an observer regulates their path to keep the target held at a fixed direction (i.e. 'cancels drift') that is away from straight ahead, the observer will reach the target by a curved path. It may seem obvious that the former, straight-ahead, trajectory would always be the simplest and most obvious to use. Below we show that this is not necessarily the case. First, let us define a few terms and show how different fixed eccentricities change the resultant trajectory.

We will define the position 'straight ahead' of an observer's body mid-line as 0 degrees and then specify ego-centric directions (measured in the cardinal plane or horizontal meridian) away from straight-ahead in degrees of visual angle. We will adopt the term 'eccentricity' to describe the ego-centric direction, or azimuth, measured in degrees of visual angle (see fig 6.1). We define this angle as α.

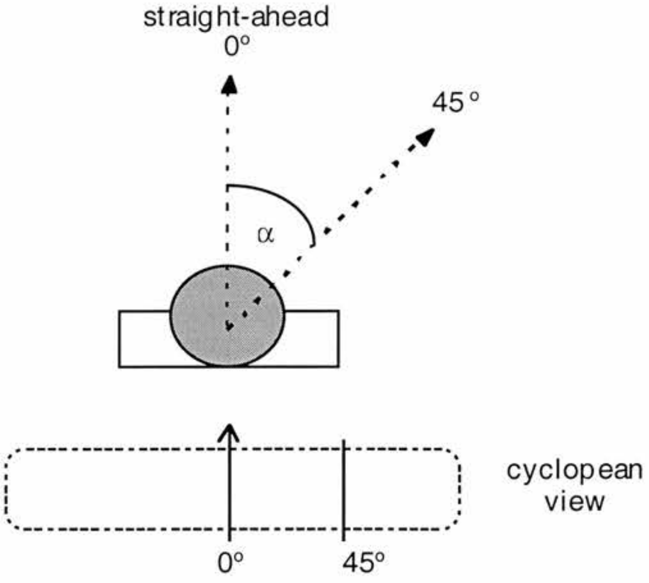


Figure 1: Ego-centric directions, 'eccentricity', α, measured angle in cardinal plane. The 'cyclopean view' at the bottom of the figure is the view as if looking through a single eye at the centre of the head, i.e. the phenomenal view.

In figure 2, below, the trajectories that result from holding a target at differing constant eccentricities are charted. The two simple iterative equations that describe an observer's trajectory are:

$$x_{n+1} = x_n + d \sin [\, \alpha - \tan^{-1}(x_n / (D-z_n)) \,] \qquad [1]$$

$$z_{n+1} = z_n + d \cos [\, \alpha - \tan^{-1} (x_n / (D-z_n)) \,] \qquad [2]$$

where, x_n and z_n are the x and z positions at time n, d is the distance moved in a single time step, α is the chosen eccentricity and D is the distance from the start-point of the observer to the target (see fig A1). The derivation of the above equations are given in the appendix.

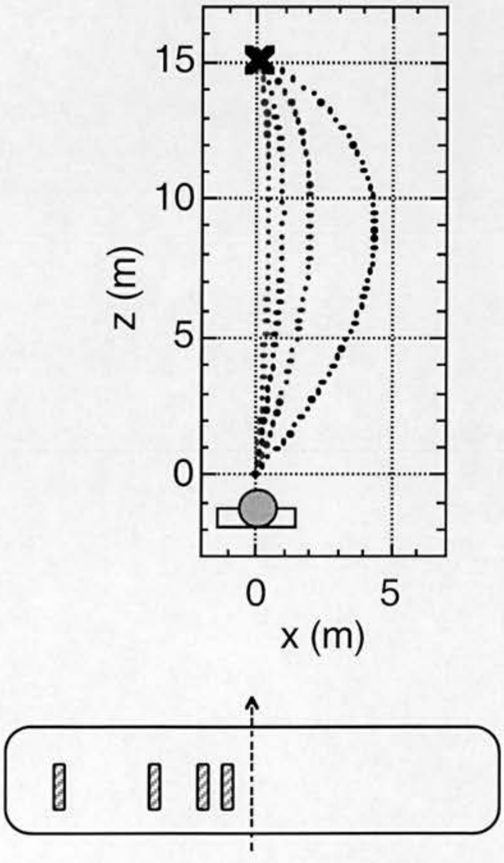


Figure 2 Trajectories (upper panel) that would result from holding a target at 5, 10, 20 and 40deg eccentricity (lower panel shows view of target as observer would see it). Holding the target 'straight ahead', i.e. at 0 deg would produce a straight trajectory leading directly to the target. Any other trajectory based upon stabilisation at an angle other than zero results in the observer 'veering' to one side before finally reaching the target.

It can be seen from figure 2 that a whole family of trajectories exist, a different one for each eccentricity, α, all of which will lead an observer to their target if they keep the target at a fixed eccentricity by 'cancelling drift'.

Equi-angular spirals and the prism trajectories

Thompson (1966) described a family of equi-angular spirals that can be found in a variety of natural forms. Do the prism trajectories described here belong to the same family of shapes? If so, why do they look different and why don't they hook round at the end?

Yes. The prism trajectories are equi-angular approaches. The reason they don't look like traditional equi-angular spirals is because descriptions of equi-angular spirals normally start with deviations from a circle (i.e. difference between instantaneous direction of movement and the control point is reduced from 90 degrees) and the prism trajectories described here start from a straight line (i.e. the difference between the instantaneous direction of movement and the control point is increased from 0 degrees), consider the following description of somebody walking towards a fixed target to see this point:

Prism/fixed eccentricity trajectories: If the person keeps the target straight ahead (0deg) then they will take a linear trajectory direct to the target. If they walk so as to keep the target at 10 deg to themselves (relative to the body mid-line) then they will take a slight curving trajectory towards the target (see section 6.2.3 for a graph of trajectories corresponding to 0, 5, 10, 20 and 40deg eccentricity).

A person walking an equi-angular spiral: If the person walks so as to keep the target at 90 deg to themselves (relative to the body mid-line) then they will walk in a circle around the target. If they now change their strategy to keep the target at 80deg to themselves then they will slowly spiral in towards the target. As the angle is reduced from 80deg to 70 deg so the spiral becomes tighter and the target is 'circled' less times. Once the angle is down to a low value the person no longer circles the target but instead takes a curving trajectory.

The same 'maintain the target at an eccentricity of X' strategy is employed in both cases.

6.2.4 Different trajectories

Let us compare two trajectories, 0° or straight-ahead, and 40°.

'Straight-ahead' trajectory: Initially the target is straight-ahead. To maintain this path it is necessary to monitor for any lateral drift of the target and correct when necessary. If the observer glances away for a few seconds, whilst maintaining their direction of motion, they are still likely to be approximately on the correct path (the target is 'stable' and not prone to rapid drift).

40 ° trajectory: To maintain this trajectory it is necessary to **change** direction slightly after every step. If the observer glances off for a few seconds and does not regulate their direction of locomotion then they will quickly find themselves a long way off course. Correcting for such a lapse will require a large compensatory change. The position of the target is thus 'unstable', it will readily drift if the change of direction is not exactly correct and continuously maintained from step to step. Because it is necessary to change direction every step (and because the rate of change of direction increases through the trajectory), the observer will be subject to centripetal forces that make it more difficult to walk. The 40 ° trajectory is also longer than the 0 ° trajectory and can only be taken at a high speed if grip and stability allow.

In summary, a large eccentricity trajectory is longer and more *'difficult'*, i.e., direction of locomotion needs to be continuously regulated and there is a greater consequence if there is a lapse in regulation. Below we quantify this.

6.2.5 Difficulty and Target Drift (ω)

Here we will start to move beyond Llewellyn's model. Following on from the above let us define what makes one trajectory more difficult than another. Consider an observer on a trajectory with the target at an eccentricity, α (fig 3 below).

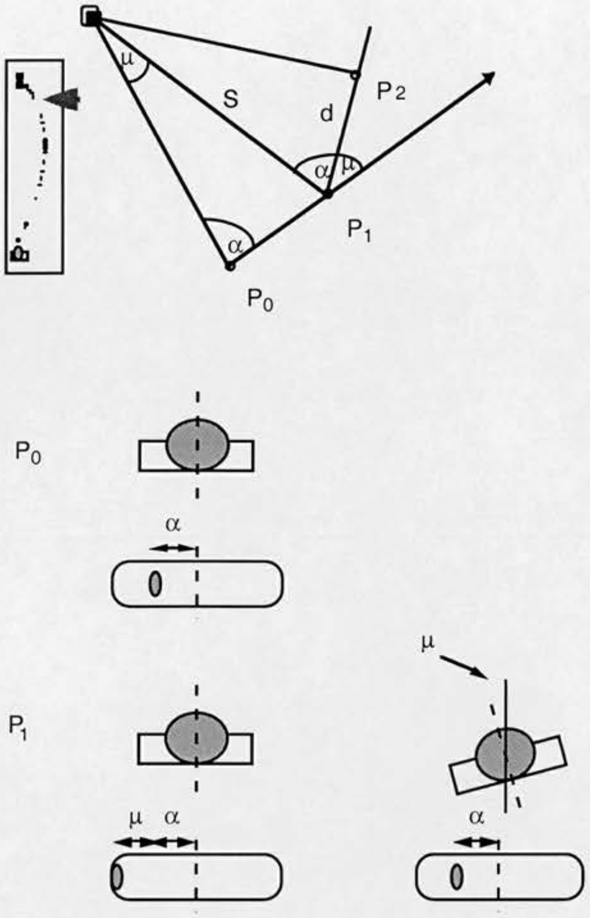


Figure 3 Top panel shows an observer walking towards a target on a fixed eccentricity, α, approach. **Inset** shows an expanded view of the position along the trajectory. At P1 the observer may either continue in the same direction, or regulate their trajectory so as to reach P2. **Lower panel** shows these two alternatives. At P1 the target has drifted by υ degrees to a new eccentricity (**left**). The observer can cancel the drift by rotating through υ degrees and putting the target back to an eccentricity of α (**right**).

The figure shows that when an observer moves forward the target drifts an angle μ, and it is necessary to regulate direction of approach (by turning through μ degrees) to return the target to the desired eccentricity. μ is dependent upon the eccentricity, α, chosen, the stride length, d, and the distance to the target, S.

For a given foot-speed, v, as μ gets larger, the trajectory gets more 'difficult', i.e. the observer needs to turn at a faster rate. However, walking the same trajectory at a slower speed reduces the rate at which the observer needs to turn and hence reduces difficulty. Thus, the critical parameter is not target drift, μ, but $d\mu/dt$, the rate of target drift, which we define as ω.

Since the speed in a direction tangential to the observers position is given by v sin (α), ω, the angular velocity is given by:

$$\omega = \sin(\alpha)\, v / S \qquad [3]$$

Where v is the speed of travel, S is the distance to the target and α is the eccentricity of the target. It can be seen that the above equation [3] incorporates each of the parameters that influence difficulty.

Figure 4 overleaf charts the relationship between the three parameters.

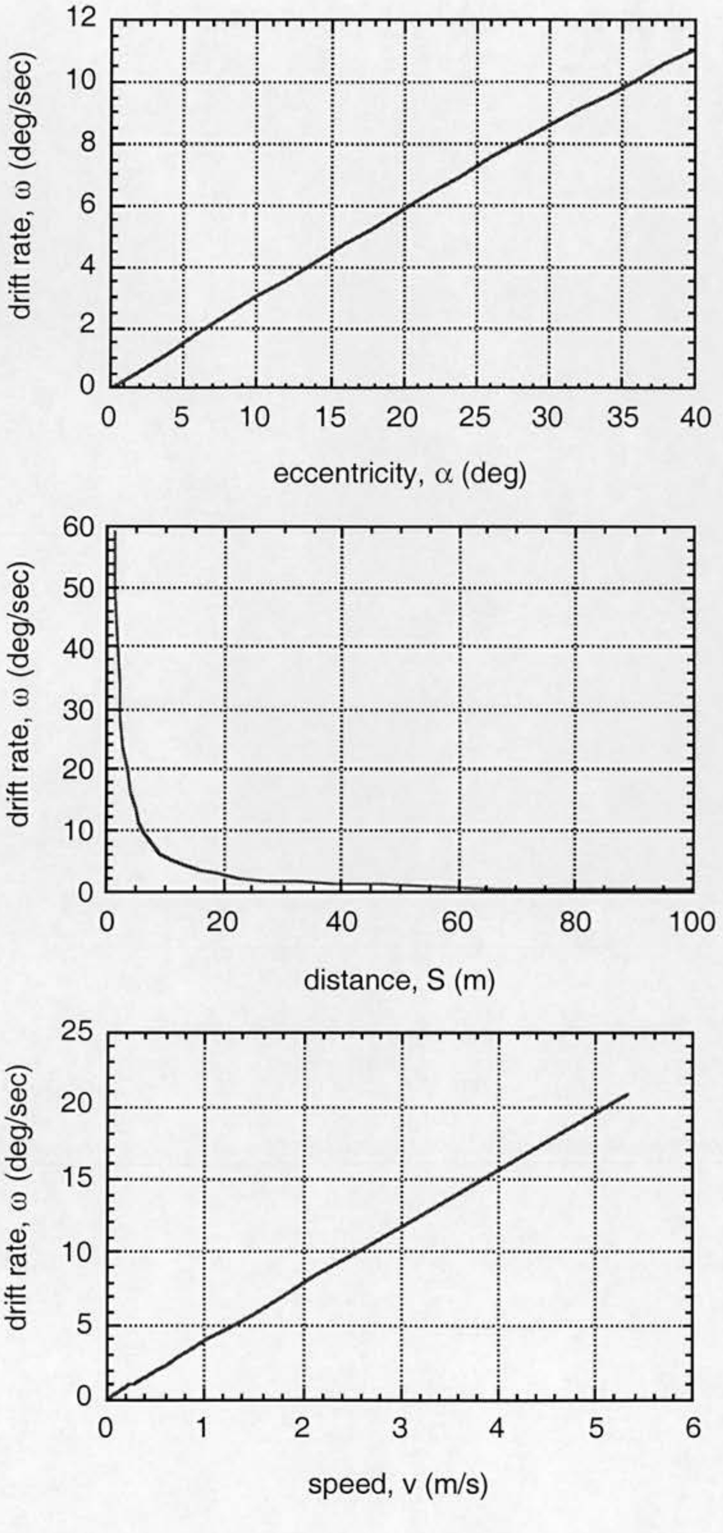


Figure 4 Rate of target drift, ω, plotted as (a) a function of eccentricity, (b) distance to the target and (c) foot speed. Values of parameters not varied are eccentricity, α = 20deg, distance, S = 5m and foot speed, v = 1.5m/s.

Note, it also follows that an observer could use an inverse model to derive their eccentricity, α, from ω. The direction that is least *difficult* and most visually stable is that which corresponds to taking a 0° straight-line trajectory to the target. Thus if an observer is moving on the most stable path, they must be on a 0° trajectory.

From figure 4 it appears that curved trajectories are an unnecessarily difficult way to reach a target. Why might observers ever take a trajectory other than a straight 0° eccentricity trajectory?

6.2.6 Non-straight trajectories

A capacity to follow non 'straight ahead' drift cancellation trajectories could be useful for three reasons: (i) final orientation and approach; (ii) general solution for approach to an object, i.e. objects off to one side as well as straight-ahead, and (iii) it may be simpler to take an approach based upon holding the current eccentricity of the target fixed, rather than translating between co-ordinate frames to find the 0° direction. Taking these points in order.

6.2.6.1 Final orientation and approach

Imagine the task of approaching a person standing directly in front of you, but who is oriented at right-angles to you, looking towards your right (see fig 5). A straight-ahead trajectory would take you directly to that person (a in fig 5). However, non straight-ahead trajectories may prove useful in circumstances when final orientation and approach are a consideration. For example, it may be desirable to (i) keep the target at a fixed eccentricity to the right of straight-ahead - this would allow you to sidle up to the person from behind unnoticed (b in fig 5) - or (ii) keep the target at a fixed eccentricity to the left which would make you visible to the person during much of your trajectory, allow you to make eye-contact during the latter part of the approach and finish almost facing the person (c in fig 5). In contrast, when walking over to stamp on an ant, the orientation of the ant relative to you does not matter!

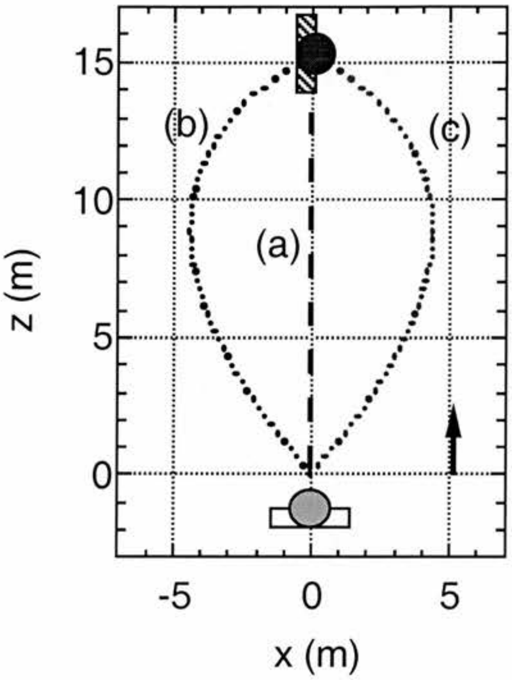


Figure 5 Approach to a target using an eccentric trajectory.

6.2.6.2 General approach solution

Holding a target at a fixed eccentricity is a general solution for reaching any target. We have already described above the approach to targets that are straight-ahead. If you approach a target that is off to one side, regulating direction of locomotion to keep the target at the same eccentricity will bring you on a smooth curve towards the target (see fig 6).

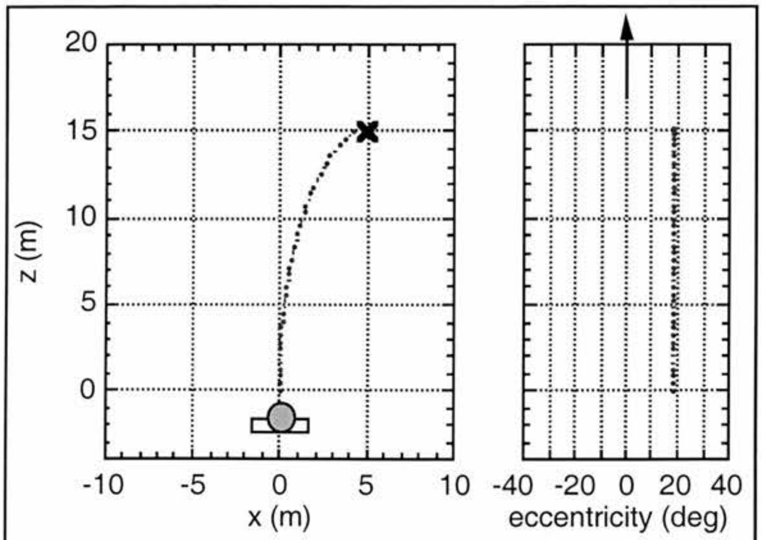


Figure 6 Approach to a target off to one side using a fixed eccentricity strategy. **Left panel** is plan view of the trajectory, **right panel** shows eccentricity, α, as a function of z. When the observer notes an object of interest that they wish to approach, all they need to do is hold the target at its current eccentricity, α (as shown in the right panel).

6.2.6.3 Co-ordinate frames

For some tasks involving the control of approach it can be difficult to establish the direction 'straight ahead'. Under some circumstances although a straight-ahead trajectory may be optimal it may not matter (i.e., time, grip or distance are not a concern). A non straight-ahead drift-cancellation trajectory will still get you to your target. Translating between co-ordinate frames to determine the straight-ahead direction may require more effort than is warranted.

This leads us into a new point. The trade-off or limits associated with eccentric trajectories.

6.2.7 Constraints on trajectory choice

Consider an example: if an observer is walking on ice (and thus has poor grip) or on high heels (poor stability) then they may find a 40 ° trajectory very difficult (or even impossible unless taken very slowly) to walk without falling over. In similar circumstances, if the target is off to one side then an observer may not wish to have to stop to 'spin on the heel' to orient themselves at the beginning or end of the trajectory. In this case, a smooth trajectory that spreads any turn over the whole course of the approach will give the observer most chance of reaching their target comfortably.

For a given speed there will be a maximum rate of change of eccentricity, ω that an observer can attain without falling over. This may constrain the choice of trajectories or require the observer to slow down when the rate of curvature increases.

Reconsider fig 4 which plots the ω as a function of eccentricity, distance and speed. Through empirical measurement it would be possible to mark limits on the abscissa for (i) comfort and (ii) stability for different surfaces, types of footwear and individuals.

6.2.8 Trade-offs between different strategies

If there is a 'comfort zone' that constrains path choice then under normal circumstances that will leave a range of responses available to an observer. An observer may trade-off advantages and disadvantages associated with different responses and so decide on one approach to use rather than another. Alternatively an observer may rather arbitrarily pick one approach through habit or random choice.

Under natural circumstances a number of factors may influence path choice and these are described in a later section.

6.2.9 Default paths

With respect to path choice we argue that a 0° path is the **default** choice as it is the easiest trajectory. It also the shortest and quickest. Unless there is reason to not take this trajectory the observer will choose it.

6.2.10 Summary so far

We have reviewed Llewellyn's drift cancellation model and described the set of trajectories that result from its use. We have introduced the notion of *difficulty* (as defined by rate of target drift, ω). We have also described some of the factors that may influence path choice.

Before we proceed to examine experimental data or discuss the use of such a strategy in natural contexts we first cover existing theories of perception and/or control of locomotor direction and discuss where they may prove useful for guiding locomotion in the natural world.

6.2.11 Motion parallax, Splay & Optic Flow

Motion parallax: when moving through an environment it is not only the eccentricity of the target object that may change. Other objects will also be seen to change eccentricity. Cutting (1986) has championed the exact opposite of the theory we have outlined so far. He suggests that the eccentricity of the target is not important but rather 'differential motion parallax' (DMP), the relative displacement of all objects in the scene (so the relative change in eccentricity of objects in the scene in our terms) is the key parameter. The relative displacements of scene objects are dependent upon the observer's direction of locomotion. So in principle, if the relative displacement of different scene objects is known, it is possible to determine the direction in which you are travelling. There is considerable evidence that this source of information can be used in determining direction of locomotion or heading in computer generated displays containing a small number of scene objects (e.g. Cutting et al, 1991).

Splay: Beall and Loomis (1996) looked at the influence of splay on performance when steering down at road or path[4]. Splay is the angle

---

[4] They also looked at 'heading-relative bearing' which is the same as eccentricity by our definition. They found that when close portions of the road were visible that 'heading-relative bearing' information alone was sufficient for optimal performance (see Beall & Loomis, 1996,

between a line perpendicular to the horizon and the edge of a road or path. It is possible that splay could prove to be a powerful influence on an observer walking a path. However, it should be noted that splay or splay rate is only useful when following an explicit path containing clear edges. Splay is therefore interesting but pertains to a limited case.

Optic Flow: Invariants in the radial spatio-temporal luminance pattern on the retina specify the locomotor direction either with respect to the retina or objects within the environment. In particular, Grindley (cited in Mollon, 1997), Gibson (1950) and Calvert (1950) noted that the 'focus of expansion', the point from which all motion appears to radiate, specifies the direction of locomotion for a fixed gaze. It has been shown that locomotor direction can be determined (by highly practised observers) with sub-degree precision from displays containing random dots distributed in depth (Warren & Hannon, 1988). It has also been shown that it is possible to accurately determine locomotor direction from sparse environments that consist of only a dozen or so objects. See Warren (1995) for a comprehensive review of these matters.

Alternative control strategies: An implicit assumption underlying much research on 'heading' is that natural *control* of locomotor direction is reliant upon the use of the percept studied in judgement research. This need not necessarily be true. Independent systems may responsible for (i) conscious perception and (ii) perceptual control. Neuropsychological research has documented many convincing demonstrations of dissociation's between perceptions and actions. Goodale & Milner (1992) have provided the most recent eloquent exposition of this position. Thus judgement tasks may have limited usefulness in describing control.

---

fig **6)**. However, two sides of the road were visible so it is not possible in their study to distinguish between the use of eccentricity and differential motion parallax (relative displacement of left and right sides of the road markers).

An example of a simple strategy that may be used for controlling direction of locomotion is equalising flow. Srinivasan, Lehrer, Kirchner & Zhang (1991) have shown that bees equalise the flow in the left and right hemifields to allow them to fly in a straight line down a corridor. Recently it has been demonstrated that humans can use similar information (Duchon & Warren, 1998).

Another strategy that would allow a target to be reached is to attempt to place the focus of expansion of the optic flow field over the target (Gibson, 1958).

Use of other simple strategies or information is quite likely as control of locomotion involves negative feedback. For example, with a feedback loop it is possible to navigate solely using cues that provide nominal travelling left-of-reference vs. travelling right-of-reference information. Therefore, it is not necessary to rely on sources of information that specify exact locomotor direction.

6.2.12 Gaze and eccentricity

Before we start considering the data that provides insight into the respective influences of DMP, eccentricity, splay and flow, we will first make a quick diversion and address a few concerns about gaze direction.

<u>The role of gaze information:</u> Heading judgement research has recently been pre-occupied with studying an observer's ability to judgement of their locomotor direction whilst making gaze movements (Warren & Hannon, 1988; van den Berg, 1993l Royden, Banks & Crowell, 1992; Cutting et al, 1992). This research was prompted by the theoretical question: is the optical flow field sufficient for perception of locomotion, or is it necessary to add 'extra-ocular' gaze direction information? This question arises from the problems that occur for a purely flow based solution when eye or head movements occur and dramatically alter the flow (Longuet-Higgins & Prazdny, 1980).

It is an important assumption of the eccentricity model that gaze-direction information is used. First, we note gaze movements do not cause a problem for a model based upon eccentricity. The use of gaze direction in our eccentricity model is very different. Gaze movements are only problematic when decomposing a flow field.

Second, we point out that eccentricity is not exclusively given by the sum of retinal position and extra-ocular gaze direction (eye position + head position). Under natural conditions eccentricity is specified optically. Shoulders, sometimes feet (sometimes stomach!) and nose all provide optical references. Therefore it is strictly only necessary to have gaze information when in a Ganzfeldt. Further for a fixed foot speed, rate of target drift, ω is a function of eccentricity (see fig 3). If the target slips more readily when it is held at a new eccentricity then the new eccentricity is further from 0°. If the eccentricity is ambiguous then rate of target drift may be a useful cue.

Precision of perceived direction: Cutting (1986) discussed the potential use of eccentricity of the target when he reviewed potential sources of information for determining direction of locomotion (specifically he was evaluating Llewellyn's drift cancellation theory). However he dismissed its potential. The reason for this was two-fold: first Llewellyn's own data can be interpreted as suggesting that observers can only manage to discern the direction of locomotion to an accuracy of about 6 degrees when presented with a display that allows use of a drift-cancellation or eccentricity strategy. Second, Cutting noted that his main interest was in the use of purely optical information. Consequently, he ruled out Llewellyn's drift cancellation theory and thus eccentricity, and unfortunately did not pursue this approach further.

We should quickly address these points now. First, as noted above, eccentricity is specified optically and so would fit with Cutting's (arbitrary) criterion for a suitable source of information about locomotor direction. Second, with different viewing conditions (i.e. with the possibility of using optical eccentricity information) it is likely that the accuracy that could be obtained in a similar judgement study would be considerably better.

Third, target drift is function of not just locomotor direction relative to a target but also distance of the target and speed of approach (see [3]). Figure 4 graphically illustrates this relationship, it shows the rate of change of eccentricity, or target drift, ω, that occurs when an observer moves towards a target object, initially at 10 ° eccentricity. It can be seen that when the target is distant it barely drifts, however, when the target is close it undergoes a large change in eccentricity (or drifts rapidly). Therefore, even a small error in locomotor direction produces a large target drift when a target is near. In other words, due to geometrical constraints, even a limited ability to judge eccentricity, α, would get an observer first to a tree and then around it without collision, because when the target is very close, any drift indicating even a small error in direction will be large. This 'improvement' in an observer's ability to judge direction of travel relative to an object as it approaches is exactly what is needed under ecological conditions and so even a poor ability to judge eccentricity, α, can be very effective in guiding locomotion.

Lastly, even if an observer has limited ability to judge absolute eccentricity, α, they are exquisitely good at perceiving a *change* in eccentricity, ω, i.e. motion. This is especially so when viewing a lit ambient environment full of reference objects against which motion can be seen. Therefore it should still be possible to instantaneously correct eccentricity by cancelling drift rather than explicitly judging eccentricity (such a strategy would probably result in a slow positional drift).

Eye-movements do not pose undue problems in a lit environment. They are more problematic in a Ganzfeldt and Freeman & Banks (1998) have recently quantified the retinal and extra-retinal determinants of position drift during pursuit eye-movements.

To summarise, we believe that there is no *a priori* reason to suppose that observers could not use eccentricity to guide locomotion.

6.2.13 Eccentricity, predictions and data

So far we have described a model of control of locomotion based upon eccentricity. We have also briefly skimmed a variety of traditional theories of perception of locomotor direction. Now we have set out the theory we can consider some data that suggests whether such an eccentricity strategy is indeed used.

The two studies we will consider both rely on manipulating the perceived location, or eccentricity, of a target and asking an observer to walk directly to it.

The first was a study we have recently reported that used optical prisms (Rushton et al, 1998a), the second had a very similar aim, but used a computer simulated visual scene (Warren & Kay, 1997). We believe that both sets of data can be accounted for by our theory. We will briefly summarise the results of both studies and then consider how they fit with any or all of the strategies described above.

6.2.13.1 The prism study

We recently reported a study that involved prism-wearing observers. The prism study (Rushton et al, 1998) was conducted in a natural environment. Observers walked over a flat slabbed area that was surrounded by buildings and trees. They were asked to walk briskly towards a target held out by an experimenter positioned about 10m to 15m away. Observers wore left or right wedge prism glasses. The glasses deflect the image and so shift the perceived location of objects approximately 15 degrees to the right or left. Note, that this manipulation does not change the relative positions of the FoE and the target (see chapter 5).

Wearing prism glasses had a dramatic effect on the trajectory taken by observers when asked to walk towards the target. Observers veered whilst attempting to walk 'straight towards' the target.

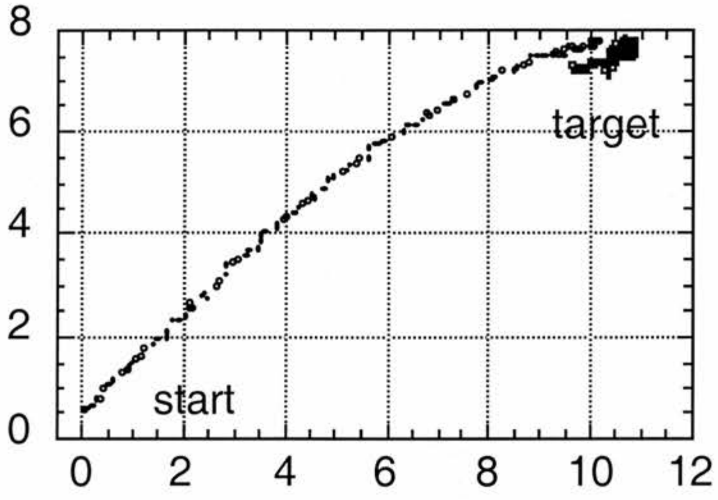


Figure 8: An example trajectory of an observer wearing prism glasses.

We described a simple model that is compatible with the data: the *perceived direction* model. The *perceived direction* model is a simple eccentricity model and it predicts that observers take a curved path if they are attempting to keep the target 'straight ahead' of them. This is explained as follows. When wearing prisms the perceived position of the whole scene, relative to the observers body, is changed by the angular deflection of the prism - so an object at 0° will be seen at approximately 15°. Keeping the target *perceptually* straight-ahead requires the observer to keep the target at a fixed eccentricity (relative to the body) of approximately 15 degrees. As was described in the theory section above, this will lead to a veering trajectory to the target. The trajectories walked by observers were very similar to those predicted by this simple *perceived-direction* model.

How does this result and the *perceived-direction* model accord with the eccentricity model? The perceived direction model is a simplified version of the eccentricity model that does not include *difficulty* nor allow for path preference. The latter concern is not relevant here as observers were given clear instructions to walk directly to the target and the surface was even with no obstacles. However, the prism manipulation dissociates the perceived 'straight-ahead' direction and minimum *difficulty* direction. A *perceived* straight-ahead trajectory is actually a $\alpha=15^0$ trajectory and the minimum difficulty trajectory is *perceived* as $\alpha=-15^0$. Therefore, we review our default straight-ahead assumption and *difficulty* and consider their likely influence.

Straight-ahead: Recall we suggested that unless there was good reason to choose another trajectory that observers would default to taking a 'straight-ahead' or 0° trajectory. However, the prisms deflect the visual world so that the direction that is perceived as straight-ahead is actually 15 ° (the angular deflection of the prism). When viewing a well lit, structured, natural environment, ego-centric directions are very well specified.

Difficulty: If an observer is going to hold the target at the position they perceive as straight-ahead then they will be walking a 15° eccentricity veering trajectory. This is a *difficult* trajectory, i.e. it requires a continuous change of instantaneous locomotor direction. However, if we consider the circumstances, the observers are young and walking on a good surface. Therefore a 15 ° eccentricity trajectory should not challenge the observer too much. However, we should recall that the trajectory will become more *difficult* towards the end.

Therefore we argue that as eccentricity, α, or ego-centric direction is well specified, and a 15° trajectory should not be too challenging for an observer to walk, we predict that an observer would endeavour to keep the target perceptually straight-ahead and so take a 15° eccentricity veering trajectory.

How does this result fit with the eccentricity theory of locomotor direction and other theories based on optic flow?

Flow: Whilst the prism manipulation alters the perceived location of all objects in the scene, it does not affect the flow (see fig 6.9). The relative positions of the flow-specified direction of locomotion and the target, measured in either image or ego-centric co-ordinates remains unchanged. Therefore perception of direction of locomotion should remain constant and veridical if flow is used.

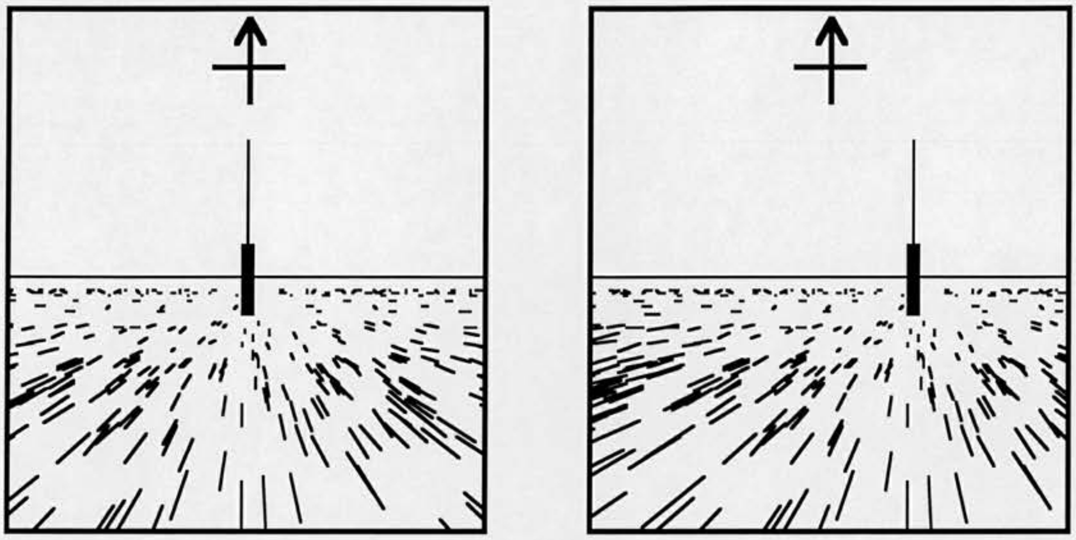


Figure 9 Forward translation (5m/s) toward tower (solid black rectangle) at 16m. Thin vertical line indicates direction of travel. Arrow indicates ego-centric straight ahead. **Left:** normal view, FoE is over tower so observer travelling directly towards tower. Tower is also 'straight ahead'. **Right**: displacement of whole image by prism. Note FoE still directly over tower indicating travelling directly towards tower. However, the tower is no longer 'straight ahead' in terms of the observers ego-centre.

Consequently perception and control of locomotion should be unaffected by the prism glasses and observers should follow a trajectory straight to the target. The flow model is therefore incompatible with the experimental results.

Differential Motion Parallax (DMP): There were plenty of trees and building features around the walking area that could have provided motion parallax information. Motion parallax would indicate that the observer is not walking straight to the target. Therefore the results are incompatible with the use of motion parallax.

An observer can try to straighten their path by explicitly using motion parallax. An informal replication (see Rushton et al, 1998) of this experiment suggests that this is possible but difficult as an observer can be seen trying to fight the natural tendency of their feet to walk in the perceived direction of the target. Introspection suggests that using motion parallax is not intuitive and requires sustained concentration.

Splay: There was not really any useful splay information present as observers walked in random directions across a slabbed area. However we informally note that observers attempting to walk down corridors (which provides rich splay information) wearing prisms walk towards the walls and do not remain in the centre of the corridor.

6.2.13.2 A closer look

Above, we described the concept of *difficulty* as described by ω, the rate at which the observer must turn, or the visual drift of the target. We also noted that ego-centric direction, α, is specified by (i) gaze direction and retinal position and (ii) optically with reference to the body. Does the prism task and dataset allow us to make and test any predictions?

Extra-retinal vs. retinal specification of eccentricity

We used two types of prisms, (i) glass wedge prisms that had a narrow field of view (FoV)and a large occluding border, (ii) Fresnel lenses that had a wide FoV but suffered a lot of optical distortion.

Field of view: from fig 10 below it can be seen that field of view may potentially have an influence on the resultant trajectory. This is because with prisms with a large FoV, optically specified eccentricity conflicts with extra-retinally specified eccentricity (lower panel). In contrast with the restricted FoV (middle panel), optical references, such as the nose, are occluded and so there is less basis for an attenuation of the prism deflection due to a conflicting optical specification of eccentricity.

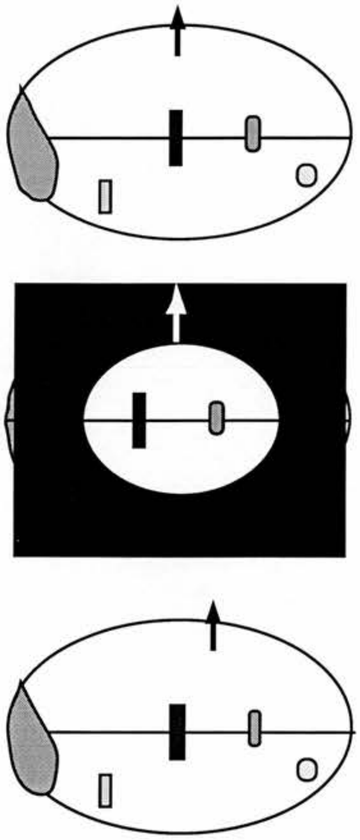


Figure 10: Monocular views of a target (solid black object towards the centre of the FoV) with nose (see Gibson, 1979) indicated in grey at the left. Vertical arrow indicates proprioceptively specified straight-ahead. **Top**, Normal view, target is straight-ahead, at 0 deg eccentricity (as indicated by arrow). **Middle,** view through a limited field of view prism. Note that target is displaced away from straight-ahead or 0 deg. **Lower**, view through a wide FoV prism. Note all of scene is displaced. Target is displaced from straight-ahead but its relation to the nose (indicated in grey) and reference objects remains unchanged. Therefore there is a conflict between eccentricity as specified proprioceptively (arrow) and optically (with reference to nose, etc.).

The conditions illustrated by the middle and lower panels are a good approximation of the wedge (narrow FoV and peripheral occlusion) and Fresnel (wide FoV) prism conditions used in the prism study.

Therefore we may expect that the wedge prisms will produce proportionately more veering than the Fresnel prisms (if optical specification of eccentricity attenuates extra-retinal specification). The figure below shows a table of the mean veering (α, eccentricity of trajectory or angle between direction of the target and direction of travel) for the two conditions. The results are compatible with this prediction.

	Mean α over trial (deg)	Angular displacement (deg) of prism	Percentage (%)
Wedge	14.4	16	90
Fresnel	8.0	14	57

Figure 10. Table of mean difference between instantaneous direction of travel and direction of the target (α) for wedge and Fresnel prisms.

We must add a note of caution and point out that the results are only compatible with the eccentricity/FoV prediction. They are also potentially compatible with a flow based explanation: the larger field of view in the lower panel of fig 10 leads to the observer being given flow information over a wider area, especially from the ground plane. We do not consider the flow explanation as likely as there should be more than adequate flow even with a narrow FoV. However, results on perception of direction with small fields of view come from laboratory simulations, so it appears prudent not to rule flow out until data has been collected under natural conditions.

The way to distinguish between these two differing explanations for the wider field of view would be to have observers perform two tasks, (i) a walking task as already described and (ii) a reaching task. If the optical specification of eccentricity with the wider FoV does attenuate the effect of the prism then there should be seen to be a reduced 'prism effect' in both the reaching task and the walking task. If the wider FoV prism shows a reduced effect when walking because of the increased amount of flow then there should not be a reduced prism effect when reaching.

Difficulty (ω) vs. straight-ahead**:** The prism manipulation pits straight-ahead against *difficulty*. The trajectory that results from keeping the target at the perceived straight-ahead position results in a curved path. This trajectory is not especially challenging for the young participants walking on the slabbed surface. However, there is a suggestion of a reduction of mean eccentricity towards the end of the time-course in the summary figure in Rushton et al (1998a). This would make sense as *difficulty* increases towards the end of the trajectory because ω increases (see fig 3). In some trials of the prism study the target was moved into or away from the path of the observer as they approached. Examination of this data in detail should provide clear evidence of the effect of *difficulty*.

When an observer starts walking they initially walk away from the target. If the target is moved into their path then the trajectory that results from keeping the target perceptually 'straight-ahead' is almost straight and so 'easy' to walk. Therefore, it is likely that the angle between the instantaneous direction of travel and the direction of the target (α), will be approximately the angular deflection of the prism.

However, if the target is moved in the direction opposite the observer's current path then the trajectory that results from keeping the target perceptually 'straight-ahead' is a lot more sharply curved. Therefore, it is possible that the trajectory will become too *difficult* (require the observer too turn too sharply) and so the observer will not be able to maintain an appropriately curved trajectory. Therefore α may be reduced in this condition.

Figure 11 below shows mean α across the time-course for trials in which the target was moving into the path and trials in which it was moving away from the observer's path. α is close to the deflection of the prism for the 'into' trails but notably reduced for the more *difficult* 'away' trials.

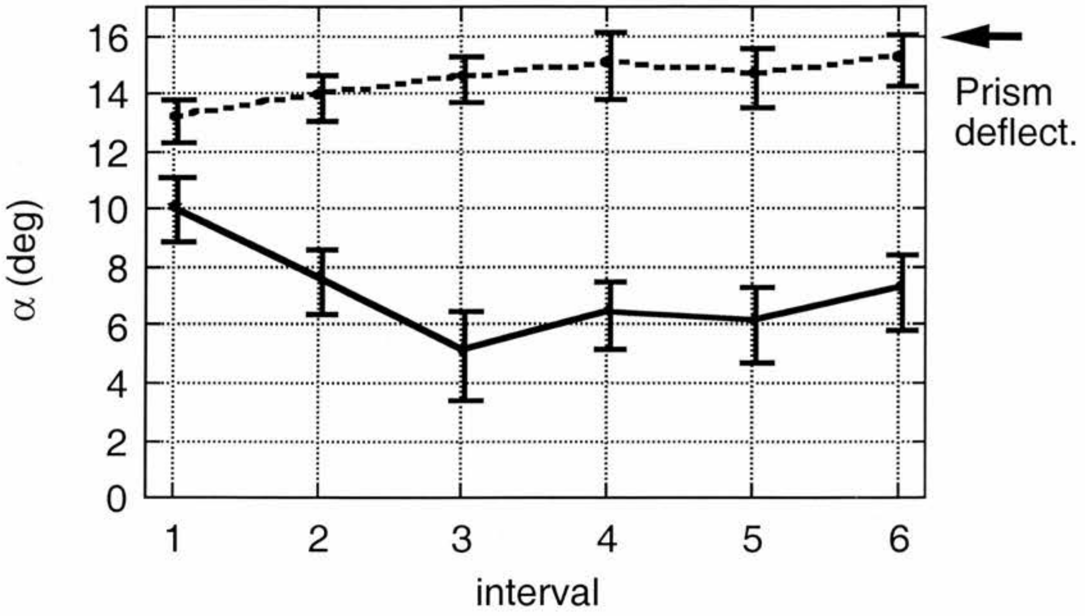


Figure 11: Trajectories for moving targets, 5 moving into path and 5 moving away from path. Length of trajectories varied so results were normalised by dividing each trial into 6 intervals. Mean α and SDs. Moving into path, dotted line. Moving away from path, full line. Note trajectories were initially smoothed by a gaussian (σ = 8 datapoints or 1/3 sec, window = 16 data points or 2/3 secs) as a consequence the first 2/3sec and last 2/3sec is lost.

This data very vividly demonstrates that observer trajectory is constrained by perceptuo-motor factors. It should be possible to reduce the eccentricity of an observers trajectory similarly by requiring them to walk faster (though there is a danger of 'overshoot') or having the observer walk in high-heeled shoes or over ice.

6.2.13.3 Simulated prisms - Warren & Kay

Warren & Kay (1997) recently reported in preliminary form some results from a simulation study. Observers walked on a wide treadmill in front of a large back-projection screen. On the screen was displayed the inside of a room with a door on the far side. The observer's task was to walk towards the door. The projection point for rendering of the display was yoked to the position of the observer's body on the treadmill (the treadmill was large enough for them to step sideways) in the plane parallel to the projection screen. Warren's manipulation was to dissociate projection direction and the direction of travel so that they were 5 degrees apart. This produces a situation similar to the prism manipulation. It has the advantage of being more manipulable than the prism study, but the disadvantage of presenting the observer with a less ecologically valid stimulus display.

In the experimental condition, when the observer started the door appeared to be 5 degrees away from straight-ahead. The observer could keep the door at that position by walking straight or they could walk to one side to bring the door across in front of them.

Warren & Kay's experimental predictions were described thus: (i) observers would move so as to place the door directly ahead of them which indicates that ego-centric direction (eccentricity) is used in the control of locomotion or (ii) observers would move so as to place the door 5 ° off to one side which would make the focus of expansion coincident with the door, indicating that optic flow guides locomotion. Warren & Kay found that observers initially place the door straight ahead and then after a second or so move so as to place the door at 5 °. This result was interpreted as supporting the hypothesis that optic flow is the dominant cue.

How would we explain this result with reference to the eccentricity theory? This study similarly dissociates perceptual straight-ahead direction and the minimum difficulty direction. Let us consider definition of straight-ahead and also *difficulty*.

<u>Straight-ahead:</u> In the experiment, straight-ahead is poorly defined, the observer is on a treadmill in front of a projection screen (a peculiar situation, the observer has to walk at a speed - specified by the treadmill else they are thrown off - but does not change position with respect to the room -- see Pelah & Barlow, 1996 for examples of odd perceptual effects). The environment is obviously 'virtual' as it is displayed at low spatial and

temporal resolution on a flat screen. Straight-ahead appears to be defined with reference to the screen edges as the environment is not fully yoked to the observer's movement (e.g. distance from screen and bounce and sway of the head). Therefore the observer may adopt an exo-centric view of straight-ahead (similar to that adopted when playing a computer game or driving a car). This is very different from walking through a natural lit environment where straight-ahead is very clearly specified.

Difficulty: Walking straight on a treadmill is fairly demanding when it is necessary to match the speed of the treadmill. Attempting to match walking speed adds an extra degree of freedom to be controlled (see previous section). It is normally hard to walk across a treadmill (it requires that the observer must walk faster so as to keep the forward component of their speed constant as they traverse the treadmill). Walking a curving trajectory across a treadmill is unlikely to be easy. Under these circumstances the *difficulty* of a curving trajectory is likely to be greatly elevated (as compared to walking through a natural environment).

Therefore under the experimental conditions used, it seems likely that the observer will be most happy walking a straight-trajectory on the treadmill (which corresponds to a 5° trajectory in the virtual environment) as it does not require the observer to more than occasionally cancel any slight drift, there is no need to continuously change instantaneous direction of locomotion and the door remains 'stable' on the display. Straight-ahead is poorly defined and the display is very obviously artificial so the observer will probably not be unduly concerned about the door being slightly away from straight-ahead. Therefore it does not appear contrived to suggest that the 5° approach taken by the observers is what we would predict from the eccentricity model. Also we believe it is more parsimonious to accept the above argument which reconciles the Rushton et al (1998a) and Warren & Kay data than to suggest that flow is sometimes used (when it is relatively poor in a simulation display rather than rich in the natural environment).

Eccentricity vs. flow: We believe that the results of Warren & Kay do not challenge the eccentricity models. However, we cannot assuage any objections to this conclusion so we present some clear suggestions about how a sceptical reader could test the flow and eccentricity theories.

A first step would be to start from an assumption that the optic flow theory is correct and establish a baseline for how well an observer can walk a curving trajectory on a treadmill. This can be done very simply in the simulation display by adding a perturbation to the observer's trajectory so that it is necessary for them to walk across the treadmill to keep the FoE on the door. This provides a reference performance level to allow evaluation of data arising from experimental conditions.

Next the dislocation between projection and locomotor direction could be varied during the trial. The flow model would predict that the observer would ignore these perturbations (and the associated swinging of the ego-centric direction of the door) as the relationship between the FoE and the door would remain unchanged by this manipulation. In contrast, the eccentricity model would predict that the observer would attempt to stabilise the door (cancel drift or maintain the door at a fixed eccentricity). The observer's ability to do so would obviously be constrained by their ability to traverse the treadmill.

Given these two datasets it should be possible to determine the respective influence of flow and eccentricity.

We believe the prism study provides strong evidence in support of the eccentricity theory. We argue that the simulated prism study reported by Warren & Kay (1997) is not incompatible with the eccentricity theory as it may first appear. Both these datasets are from tasks specially designed to remove complicating factors such as obstacles etc. Below we discuss control of locomotion in natural circumstances.

6.2.14 Ecological contexts

Thus far we have outlined a very simple model of locomotor control based upon eccentricity. We have compared it against alternative models and looked at data from two carefully designed experiments.

However, the model outlined makes a lot of naive assumptions. It disregards influences on the observer's path such as obstacle avoidance, stability etc. Also it does not allow for the over-learning of actions that are common to skilful behaviour. For instance, the complete trajectory to take an observer around a lamp post can be executed ballistically, i.e. an appropriate sequence of steps can be taken without visual feedback. Also it is very likely that the reader can walk around their house with their eyes closed. In both cases repeated practice allows a complete action to be executed open-loop, without visual feedback or guidance.

A theory that ignores other influences on observer's path choice and over-learning and open-loop control cannot be complete in a natural context. Below we describe deviations from fixed eccentricity trajectories. Discussion of the points below is not intended to undermine the simple formulation of our eccentric theory of locomotion. Rather we wish to remind the reader of the natural context in which locomotion occurs and to prompt consideration of the control laws that predict and describe the selection of the trajectories described below.

6.2.14.1 Deviations from fixed eccentricity trajectories and different strategies

Course deviation & obstacle avoidance: A standard fixed eccentricity approach takes no account of obstacles or course preference. Very often an observer does not move through a completely open environment and they may be interested in more than simply reaching their target. Collision with an obstacle is specified by the obstacle maintaining a fixed eccentricity during the observer's approach to their target. If an observer notes this he or she may modify their trajectory mid-course. This may be by avoiding the object and then starting a new trajectory or by deviating from the trajectory long enough to avoid the obstacle.

Degrees of freedom: During an approach to an object an observer invariably has to regulate not only their direction but also their speed - controlling their time to contact so as to come to a smooth halt. Thus it is possible that they will reduce the number of degrees of freedom (Vereijken, van Emmerik, Whiting & Newell, 1992) that they need to simultaneously regulate by prospectively changing their trajectory so that it need not be altered during the last few seconds. This may lead them to choose a trajectory that allows most of the change of direction to occur early on, leaving fewer variables to be controlled at the critical stage of the final approach. Alternatively it may lead to a late deviation to the course to straighten up the trajectory prior to final approach. Other cases where it may be desirable to reduce degrees of freedom are when the observer is intending to perform other actions during the course of the trajectory. For instance if the observer wishes to throw a ball during the course of the trajectory, one course of action that would reduce the complexity of the situation would be to straighten their trajectory immediately prior to launching the ball.

Let us consider some specific examples of trajectories that are derived from a fixed eccentricity strategy. We will take the case of an observer who is walking forward and suddenly takes interest in a target off to one side. (We choose this general case rather than the specific case of a target directly ahead of the observer.) The observer could immediately employ a strategy of keeping the target at its current eccentricity, for reasons discussed above. Below are some of the systematic deviations from an immediate fixed eccentricity trajectory. Figures show both the path of the observer (left panel) and the eccentricity of the target (right panel) throughout the trial.

Delayed turn: The observer may not turn as soon as their target is spotted but rather prefer to continue their current trajectory for a little longer and make the change of direction towards the target later.

A delayed turn may still rely on a fixed eccentricity trajectory strategy to approach the target, all that is changed is the time at which the turn begins (see fig 12).

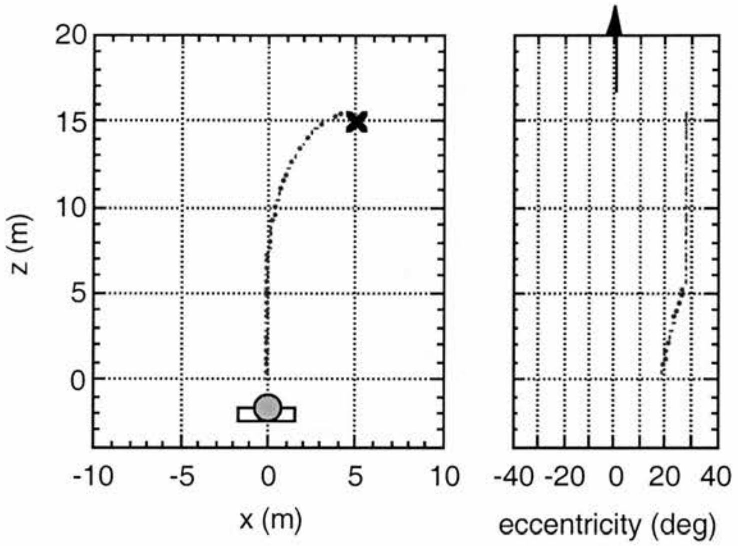


Figure 12 a late turn

Non-constant turn: An observer may not spread the turn over the whole trajectory but choose to complete most of the change of direction at the beginning or end.

They may take a sharp-early turn (see figure 13), in which they initially start to bring the target towards the straight-ahead position and then require less curvature towards the end.

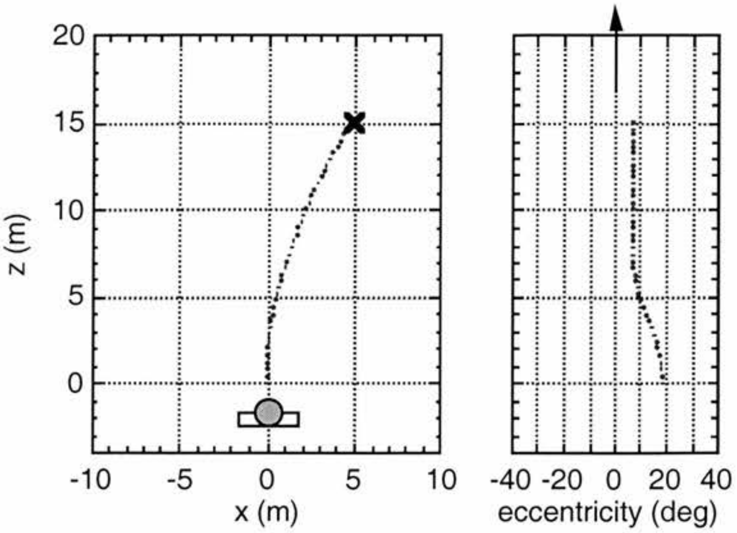


Figure 13 fast-early turn

Alternatively they may take a sharp-late turn. In this case they do begin to turn towards the target but not enough to keep the target at a fixed eccentricity. This will then require them to turn more rapidly towards the end.

Irregular turn: The observer may initially turn away from the target before starting a fixed eccentricity trajectory towards it. The result of this will be to increase the curvature of the turn that they take and hence the orientation and path during approach and arrival state (fig 14).

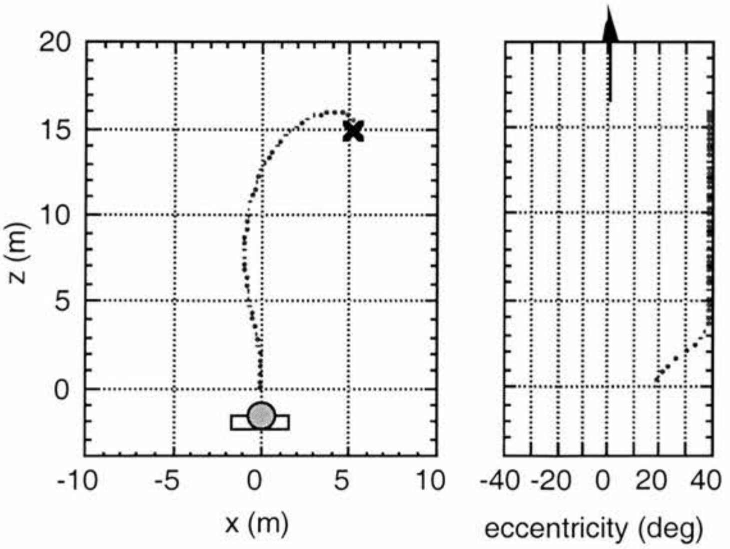


Figure 14 stepping away first.

The observer may instead turn away from or towards the target before the end of the approach. This will serve to bring them into the vicinity of the target but change their final position and orientation.

Lastly they may control eccentricity throughout the trajectory, systematically reducing it so that the target is straight ahead at the end (fig 15).

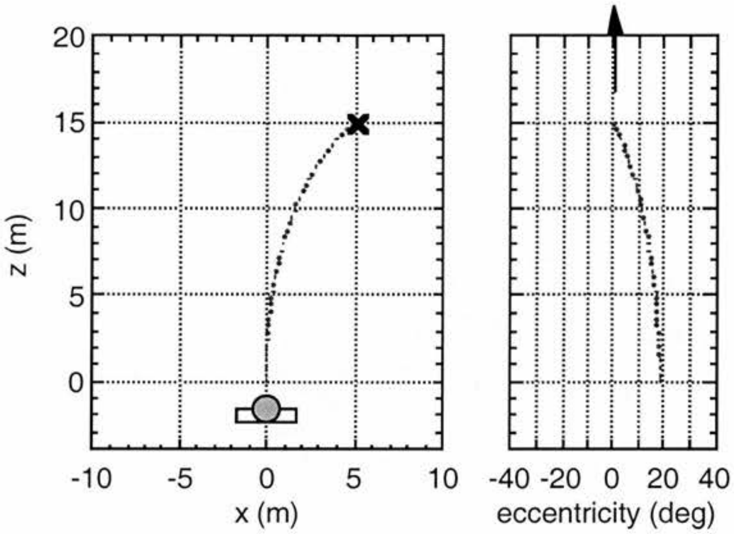


Figure 15 systematically reducing eccentricity over trajectory

6.2.14.2 Path following

We have only briefly touched upon this matter in this paper. Land & Lee (1994) and others have described the usefulness of the 'tangent point' when driving. It is proposed that control of locomotion is achieved by regulating the direction of the instantaneously-defined tangent point. An eccentricity account would require fixed or instantaneously-specified control points. For a straight path this could be the end of the convergence of the path edges or a point on the edge a fixed distance ahead. For a curved path tangent points may serve as control points. We informally note that walking along a corridor wearing prisms leads an observer into a wall. A corridor provides rich, veridical splay information. Therefore it appears that eccentricity may be the critical influence on path following.

6.2.15 Summary

We have presented a model of the control of locomotor direction that does not rely on optical flow. We have discussed the data that has been reported that bears on the model and concluded that 2 studies, one of which at first appears to contradict our theory, provide evidence to support it. We have also set out a series of predictions and ways of assessing the model.

Optic Flow, Splay and Differential Motion parallax have previously been shown to be informative under some circumstances, primarily when location information is not present. Whether they have an influence on the walking or running observer in the real environment has yet to be seen. However it should now be possible to identify when or if they do.

The eccentricity model describes the control of locomotion through regulation of eccentricity of a target. The model can be trivially extended to accommodate cross-modal influences on perception and control of locomotor direction. A high pitched noise emanating from the target should attenuate the mis-location of target eccentricity that results from use of prisms. Likewise, vestibular stimulation that modifies perception of straight-ahead should result in changes in locomotion. Another paper [Chapter 5: Afterword] discusses the changes in perception of straight-ahead that result from brain-injury and how this may influence walking trajectories (Rushton et al, 1998a).

Only further empirical investigation will clearly delineate the influence of optical flow and eccentricity and tell if the eccentricity model is to be found wanting.

6.2.17 Appendix

6.2.17.1 Derivation of eccentric paths

The observer starts at position (0,0), with the target a distance D away (see figure 1). The strategy is very simple:

(a) The observer rotates their body so that the target sits at the appropriate eccentricity, α. (An eccentricity of 0 would point their mid-line, or 'locomotor axis' at the target.).
(b) Move forward a fixed distance, d, to point (x1,y1). The target will no longer be at angle α.
(c) Repeat again from step a.

The observer's progress along the trajectory can be calculated. At an arbitrary timestep, n (see Figure 1), the position of the observer, (x_{n+1}, y_{n+1}) is calculated:

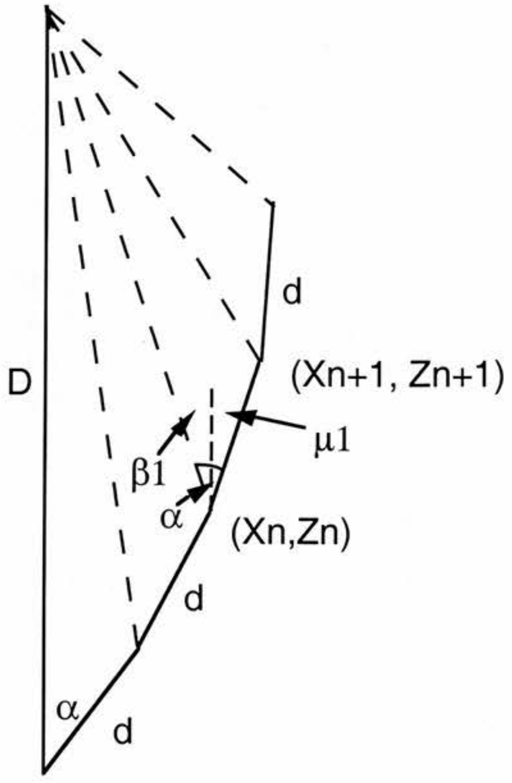


Figure A1: timesteps n and n+1 during the trajectory.

from geometry:

$$x_{n+1} = x_n + d\sin(\gamma_n) \tag{A1}$$

where γ_n is the angle made between the y-axis and the direction at time n, x_n is the previous x position, and d is the distance moved in a single time step.

and:

$$\gamma_n = \alpha - \beta_n \tag{A2}$$

where α is the chosen eccentricity, and βn is the angle between the y-axis, the observer at time n and the target.

from geometry:

$$\tan(\beta_n) = x_n / (D-z_n) \tag{A3}$$

where D is the distance from the start-point of the observer to the target.

substituting from (A2) and (A3) into (A1) gives the x-position x_{n+1}:

$$x_{n+1} = x_n + d \sin [\alpha - \tan^{-1}(x_n / (D-z_n))] \tag{A4}$$

similarly:

$$z_{n+1} = z_n + d \cos [\alpha - \tan^{-1}(x_n / (D-z_n))] \tag{A5}$$

The pair of simple iterative equations (A4) and (A5) describe the observers trajectory. The family of trajectories corresponding to different eccentricities are shown in fig 2 at the beginning of chapter 6.

References

Beall, A.C. & Loomis, J.M. (1996). Visual control of steering without course information. *Perception*, 25, 481-494.

Bruce V, Green PR, Georgeson MA. (1996). *Visual Perception: Physiology, Psychology & Ecology.* Psychology Press, Hove, UK

Calvert E.S. (1950). Visual aids for landing in bad visibility, with particular reference to the transition from instrument to visual flight. *Transactions of the Illuminating Engineering Society, London* , 15, 183-219.

Cutting, J.E. (1986). *Perception with an eye for motion.* MIT Press, Cambridge, Massachusetts, USA.

Cutting JE, Springer K, Baren PA & Johnson SH. (1991). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: Human Perception & Performance*, 121, 41-72.

Duchon, A.P. & Warren, W.H. (1998). Interaction of two strategies for controlling locomotion [ARVO Abstract]. *Investigative Optics and Visual Science*, 39(4), S892, Abstract no 4122.

Freeman, T.C.A & Banks M.S (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research*. 38, 941-945.

Gibson J.J. (1950). *The perception of the visual world.* Houghton Mifflin, Boston.

Gibson J.J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, 19, 182-194.

Gibson J.J. (1979). *The ecological approach to visual perception.* Lawrence Erlbaum Associates Inc, Hillsdale, New Jersey.

Goodale, M.A. & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in* Neuroscience., 15, 20-25

Johnston, I.R., White, G.R. & Cumming, R.W. (1973). The role of optical expansion patterns in locomotor control. *American Journal of Psychology*, 86, 311-324.

Land, M.F. & Lee, D.N. (1994). Where we look when we steer. *Nature,* 369, 742-744.

Llewellyn KR. (1971). Visual guidance of locomotion. *Journal of Experimental Psychology* , 91, 245-261.

Longuet-Higgins HC,. Prazdny K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B* , 208: 385-397.

Mollon, J. (1997). "....on the basis of velocity cues alone": some perceptual themes. *Quarterly Journal of Experimental Psychology,* 50A, 859-878.

Pelah, A. & Barlow, H.B. 1996. Visual illusion from running. *Nature,* 381,283

Royden CS, Banks MS, Crowell JA. (1992). The perception of heading during eye-movement. *Nature,* 360, 583-585.

Regan, D.M. & Beverley, K.I. (1982). How do we avoid confounding the direction we are looking with the direction we are moving? *Science,* 215, 194-196.

Rushton, S.K., Harris J.M., Lloyd M.L. & Wann J.P. (1998a in-press). The role of perceived location during locomotion on foot. *(Current Biology)*

Rushton, S.K., Harris, J.M & Wann, J.P. (under-review). Steering, Optic Flow and the respective importance of depth and retinal motion distribution.

Srinivasan M.V., Lehrer M., Kirchner W.H. & Zhang S.W. (1991). Range Perception Through Apparent Image Speed In Freely Flying Honeybees. *Visual Neuroscience,* 6, 519-535.

van den Berg AV. (1993). Perception of Heading. *Nature,* 365: 497-498.

Vereijken, B., van Emmerik, R.E.A., Whiting, H.T.A. & Newell, K.M. (1992). Free(z)Ing Degrees Of Freedom In Skill Acquisition. *Journal Of Motor Behavior,* 24, 133-142.

Warren, W.H. (1995). Self-Motion: Visual Perception and Visual Control. IN Epstein, W. & Rogers, S. (eds). *Perception of Space and Motion* , Academic Press, San Diego, 263-312.

Warren WH, Hannon DJ. (1988). Direction of self-motion is perceived from optical-flow. *Nature,* 336, 162-163.

Warren, W.H. & Kay, B.A. (1997). The focus of expansion is used to control walking. IN Schmuckler, M.A. & Kennedy, J.M. (eds). *Studies in Perception and Action IV*. Lawrence Erlbaum Associates Inc, 207-210.

Chapter 7:
Discussion and Conclusions

7.1 Perception of Time-To-Contact

In chapter 2, I presented a model, the *dipole* model, to account for the responses of observers in a ball catch task. The model includes the early combination of disparity and looming. The model was tested by a "cue-plateau" manipulation. The model demonstrates a number of features: it can deal with cue drop-out and biases its response towards the cue that specifies the earliest arrival; the weighting applied to size and disparity changes according to the size of the object; it is a very simple model; it is compatible with previous psychophysical findings.

The *dipole* model combines size and disparity very early and is a "strong fusion" model. If the model is correct then it should not be possible to change the weightings of size and disparity. Chapter 3 reported the results of a task that was designed to test this hypothesis. The results betrayed no indication of a reduction of weighting of disparity.

7.2 Perceptual control of locomotor direction

The results presented in chapter 4 suggest that perceptual control of locomotor direction is not aided by depth information. Neither disparity defined depth nor a pictorial cue to depth improved performance. Retinal motion distribution was found to be an important determinant of performance.

In chapter 5, I explored the role of perceived location. Trajectories of the observers walking wearing prism glasses were recorded and analysed. The results suggest that perceived location rather than retinal or optic flow guides locomotion on foot.

This finding undermines a widely held assumption about the central role of optic flow. The study was concerned with locomotion on foot. It remains an open question whether perceived location is used when driving a car or flying a plane.

The conclusion that flow is not used when walking but can be used when ego-centric direction information is not available raises some interesting questions: Why can we use flow? Might the ability be due to modern exposure to films and television? Might flow have an alternative use, for instance if a rock is heading towards your head there will be a focus of expansion. The position and change of position of the focus of expansion will tell you if the rock is going to hit you between the eyes[5]..

In chapter 6 a fuller model of control of locomotor direction based upon perceived location was advanced. The model appears compatible with the available data. It remains to be seen if it stands up to further testing.

7.3 The "input" to modular systems

This thesis began with a discussion of modular systems. A hypothesis was put forward that modular systems may rely on the minimal set of information that specifies a property or relation. Interceptive timing and perceptual control of locomotion, two favourites of direct perception, were chosen as the best candidates for a modular implementation that could be examined.

---

[5] However, this example raises another question: rather than asking whether flow may have a use in determining the direction of travel of an approaching object, maybe we should be asking if perceived location influences perception of direction of travel for a ball thrown towards us?

The results on interceptive timing are awkward. Both looming and disparity are used. We can explain this away by pointing out that neither disparity nor looming individually specify TTC to an observer with a limited sensitivity visual system over a full range of natural conditions. Additionally, we can collapse the distinction between disparity and looming with a new classification, the "dipole".

The results from the perceptual control of locomotion are also problematic. No sooner had data been collected that argued for a single source of information, the velocity flow field (without depth), than it became apparent that the basic skill of walking does not use this source of information. A single source of information, perceived location supports guidance of locomotion on foot. There is no indication that it is used in combination with optic flow.

It is probably fair to conclude that the hypothesis about information sources survives intact but it has taken a lot of battering along the way.