Journal of Experimental Psychology:
Human Perception and Performance
2003, Vol. 29, No. 2, 363-378

Controlling Steering and Judging Heading: Retinal Flow, Visual Direction, and Extraretinal Information

Richard Wilkie and John Wann
University of Reading

The contribution of retinal flow (RF), extra retinal (ER), and egocentric visual direction (VD) information in locomotor control was explored. First, the recovery of heading from RF was examined when ER information was manipulated; results confirmed that ER signals affect heading judgments. Then the task was translated to steering curved paths and the availability and veracity of VD was manipulated with either degraded or systematically biased RF. Large steering errors resulted from selective manipulation of RF and VD, providing strong evidence for the combination of RF, ER, and VD. The relative weighting applied to RF and VD was estimated. A point-attractor model is proposed that combines redundant sources of information for robust locomotor control with flexible trajectory planning through active gaze.

We routinely locomote through the world and reach our goals successfully despite course changes and variable eye or body positions. This requires a highly efficient system for processing visual information to pick a path, maintain a course, and arrive at our desired destination. Solutions to these problems have centered on the use of optic flow and the detection of locomotor heading. Gibson (1958) proposed that to aim locomotion at a target object, the observer should "keep the center of flow of the optic array as close as possible to the form which the object projects" (p. 187). This is now recognized as a very limited solution for two reasons: First, Gibson was addressing the case of a linear locomotor path, which produces a pattern of optic flow expanding radially with a clear focus of expansion (FoE; see Figure 1A). In the case of radial flow the FoE is coincident with the locomotor heading and the locomotor path is indicated by vertical vectors in the optic flow field. In the case of a curved path there is no FoE for optic flow, but it may be possible to discern "locomotor flow lines" (Lee & Lishman, 1977; see Figure 1C). A simple generalization of Gibson, therefore, would be to steer by keeping the target object centered within the locomotor flow lines, irrespective of whether they are radial (linear path) or curved. A critical problem is that both Gibson and Lee and Lishman were referring to *optic flow*, which is the pattern of flow that would be projected on an image plane that is fixed with respect to the locomotor axis (e.g., a car windshield). The pattern of flow on the retinal image is affected by gaze motion (Regan & Beverley, 1982). If observers stabilize gaze with respect to their vehicle, they should indeed be able to perceive locomotor flow lines. However, if they adopt the more natural behavior of fixating features within the environment, such as a target object or road features, then gaze rotation transforms the retinal flow (RF) pattern and displaces the FoE from the direction of heading (Regan & Beverley, 1982; Warren & Hannon, 1990; see also Figure

---

Correspondence concerning this article should be addressed to either Richard Wilkie or John Wann, Department of Psychology, University of Reading, Earley Gate, Whiteknights, RG6 6AL, United Kingdom. E-mail: r.m.wilkie@rdg.ac.uk or j.p.wann@rdg.ac.uk

1D). A considerable body of research has been directed at how the observer may recover heading from RF and whether extraretinal (ER) information specifying gaze motion is required (Banks, Ehrlich, Backus, & Crowell, 1996; Crowell, Banks, Shenoy, & Andersen, 1998; Lappe, Bremmer, & van den Berg, 1999; Li & Warren, 2000; Royden, Crowell, & Banks, 1994; van den Berg, 1993; Warren & Hannon, 1990). Considerably less research has been directed toward exploring the use of RF and ER information in locomotor steering (Fajen & Warren, 2000; Rushton, Harris, Lloyd, & Wann, 1998; Rushton & Salvucci, 2001; Wann & Land, 2000, 2001; Warren, Kay, Zosh, Duchon, & Sahuc, 2001). In this article we address the use of flow and ER information in both judgments of heading and active steering.

Heading and ER Information: Gaze Sweep vs. Fixation

If an observer is moving on a linear path with gaze stable, then the RF field is radial (Figure 1A). If the observer executes a gaze sweep during forward locomotion, then a rotation is introduced into the RF field. The combination of a radial expansion and a rotation around the observer results in a displacement of the flow lines, and the resulting pattern (Figure 1B) may look very similar to what would arise from locomotion on a curved path with gaze fixed (Figure 1C). The two RF patterns have very similar statistical properties, and it is plausible that even nominal knowledge that “my eyes and head are rotating to the right” may assist in dissociating the two fields. There is evidence that recovery of heading in during a gaze sweep (Figure 1B) may require ER information (Banks et al., 1996; Crowell et al., 1998), but there is also strong evidence that heading can be recovered without this information in simulations of fixation on a stable object in the environment (Li & Warren, 2000; van den Berg, 1993, 1996). This may in part be due to the reduced ambiguity in the retinal pattern: If the observer fixates a stable feature in the environment, such as a point on the ground, then the rate of gaze rotation varies over time as a function of the distance from that point. This maintains that feature both stable and centered on the retina (a necessary condition for fixation). Hence, there is a singularity in the RF, centered on the fovea, and the global flow across left and right peripheral fields is in opposite directions (Figure 1D). Although there are some complex curved paths that would produce this retinal pattern over a short time interval, the simplest, most parsimonious interpretation of the retinal pattern is of a fixation plus rotation, and we would argue that Figure 1D is not ambiguous.[1]

Our concern was with the case of ground fixation. In locomoting through the world we often fixate a location that we wish to steer toward or around. Gaze behavior generally alternates between being stabilized to the locomotor axis and fixation on the environment. In contrast we argue that we seldom sweep gaze across the ground plane, except for a transient (saccadic) motion to achieve fixation. The evidence regarding the use of ER information during ground fixation is mixed, although the results from gaze-sweep and ground-fixation experiments have often been conflated.

Two studies set the current agenda: Warren and Hannon (1988) presented their observers with two types of stimuli that simulated ground fixation during locomotion. In the first, real rotation (RR) condition, observers made eye movements to track a ground point as it moved to the side of their path before making a heading judgment. This added rotation to the scene, but it also supplied ER information. In the second, simulated rotation (SR), no eye movements were made, but the rendered scene was rotated as if the eye had moved to focus on the ground point.

---

[1] The pattern shown in Figure 1D could be produced by a downward curving path component, but the paths would have to be of changing curvature to maintain this flow pattern. This would create a new singularity to the right of the visual field, so to center the singularity, gaze would also have to be offset from the tangent to be coincident with this (unspecified) location. This highly constrained geometric solution does not render Figure 1D as “ambiguous,” which requires a reasonable degree of probability of the two explanations. All humans (and animals) continually experience the equivalent of Figure 1D as a result of fixation during locomotion; few ever experience the alternative solution.

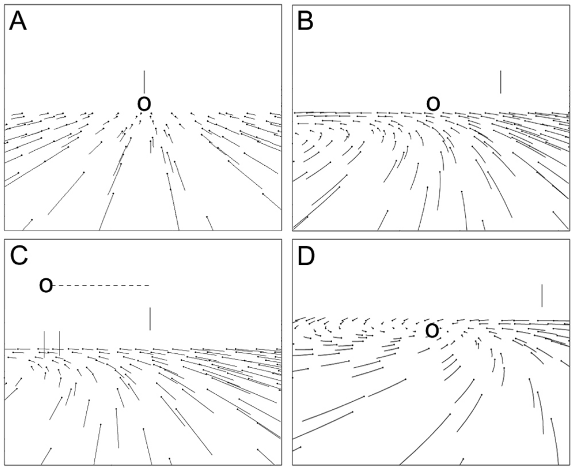

Figure 1. Retinal flow patterns arising from different locomotor trajectories and gaze responses. Panels illustrate the flow arising over 0.2 s while travelling at 15 mph (or 0.1 s at 30 mph). A: Radial flow arising from moving along a linear trajectory toward the vertical marker with stable gaze. B**:** Moving on a linear trajectory while sweeping gaze to the left. The fixation point (circle) stays centered whereas locomotor heading (vertical bar) sweeps to the right. The gaze motion introduces curvature into the flow field. Although some ground points exhibit transient vertical motion, there is no sustained singularity at the point of gaze. C: Travelling on a curved trajectory that would carry the observer through the vertical posts, with gaze locked in a forward position. Heading and gaze eventually sweep around to be coincident with the posts. D: Moving on a linear trajectory while fixating a ground feature (circle) to the left of the trajectory. Heading sweeps to the right and there is a singularity and stationary ground feature at the point of gaze.

The two conditions therefore produced the same retinal information, but crucially, the SR condition lacked any ER eye movement information. Warren and Hannon reported that observers could discriminate headings equally well in both conditions, and therefore they concluded that ER information is not necessary for the judgment of heading. The eye movement velocities used were slow (<1º/s), equivalent to walking while targeting a distant object. Royden et al. (1994) demonstrated that, for rotation rates up to 5º/s, retinal information was not sufficient to judge heading accurately, whereas additional ER information did allow correct heading judgments to be made. This seemed to support an alternative model of heading perception, where signals from the extraocular muscles of the eye are used to subtract the eye rotation from the RF, providing the observer with a correct estimate of heading. The tasks used, however, required gaze sweep and fixation within a cloud, both of which provide more ambiguous flow fields than ground fixation. In addition, the display size used may have impeded accurate retinal analysis (Banks et al., 1996).

Banks et al. (1996) re-evaluated the two arguments. Using the procedure of Warren and Hannon (1988), rotation was then added to the retinal image in one of two ways: either through real eye movements (RR) or by SR, and the proportion of each type of rotation was then varied. For each of five conditions, as the proportion of RR reduced (100%, 75%, 50%, 25%, 0%) the amount of rotation due to SR increased correspondingly (0%, 25%, 50%, 75%, 100%). In principle the sum of RR and SR remained constant, and hence the RF field should have been equivalent across conditions.

Banks et al. (1996) found that the errors in judging heading were proportional to the amount of SR within the display and that three of the four experimenters performed better when ER signals were present. The results, however, were less clear when the task required ground fixation with the advantage of ER information varying across the participants. A problem with the Banks et al. study was that if the gaze motion is imperfect and there is a transient stall in pursuit, then during that transient the 100% RR condition presents a veridical, radial flow field to the retina. The other conditions (75% –> 0% RR) have increasing bias (SR) simulated in the screen display, and the pattern of errors across the conditions could be predicted from a hypothesis of disrupted pursuit. Banks et al. did not record eye movements, so they cannot discount this hypothesis. Ehrlich et al. (1998) included the recording of gaze movements for 2 participants to show that accurate pursuit was possible, but the heading errors during these trials were not reported and gaze tracking was not used during the main experiments.

A second issue with RR displays is that the fixation target sweeps toward the display edge, which provides a visual reference for the direction of motion, given that heading always lies within the sector that subtends the largest angle.[2] Royden et al. (1994) tested the role of this visual direction (VD) cue with 2 participants using a moveable clipping frame and found an advantage of VD for one participant (TRC) but not the other (MSB). This was present in Banks et al.'s (1996) Experiment 3but was removed in Experiment 4. Because the salience of the edge cue is increased when using a textured ground plane, we used a procedure that controlled for this VD cue.

A final issue with the experiments of Banks et al. (1996) is that the results arose from testing three non-naïve participants (the experimenters) with a small range of headings (–4, 0, +4°), and no statistical appraisal was made of the error pattern. We wished to reconfirm the role of ER signals with a larger group of naïve participants, with a wider range of headings, prior to translating the paradigm to a steering task.

Steering a Path

Although considerable research has been conducted into the perception of linear heading, its role in locomotor control is debatable. Wann and Land (2000) argued that most steering tasks can be completed by using RF or ER information, but without having to recover heading per se (for follow-up debate, see Fajen & Warren, 2000; Harris, 2001; Rushton & Salvucci, 2001; Wann & Land, 2001; Wann, Swapp, & Rushton, 2000). The need to judge heading while looking at a peripheral object (the Warren–Banks task) can be negated by intermittently switching gaze back to the assumed path. For linear paths, the respective roles of egocentric VD (the angle the target subtends to one's body or vehicle) and RF have been explored by asking participants to walk linear paths while wearing prisms, which displace VD (Rushton et al., 1998; Warren et al., 2001; Wood, Harvey, Young, Beedie, & Wilson, 2000). The current conclusion seems to be that both sources are used depending on their relative strength. In the prism paradigm the Warren et al. article schematically depicts the FoE as displaced from the target by the prism, but this would not be the case. As soon as participants fixated the target, the FoE would be centered on this and the flow would be weakly curved as in Figure 1D. The simplest solution to the task is to notice that the flow is curvilinear and adjust walking direction to make it radial. The prism paradigm, therefore, addresses the use of RF, but it makes no assessment of whether heading is recovered from flow.

For curved paths, *heading*, defined as the current direction of travel, is the tangent to the path and is constantly changing. If heading is recovered during curved trajectories, there is still the need for a mechanism to ensure that its rate of change is sufficient to meet the steering requirement (e.g., Fajen, 2001). There is some evidence that observers can make heading judgments while on curved paths, but it is not clear how these relate to active steering control (Stone & Perrone, 1997; Warren, Mestre, Blackwell, & Morris, 1991). The alternatives that bypass the perception of *heading* are simple heuristics that arise from the use of active gaze (looking to where you wish to steer) and the resulting pattern of RF (Kim & Turvey, 1999; Wann & Land, 2000; Wann & Swapp, 2000) or VD information (Land & Lee, 1994; Rushton & Salvucci, 2001; Wann & Land, 2000). Effective steering of a curved path could be achieved by using information from gaze angle and gaze velocity alone, or by using the RF alone, without recovering heading (Kim & Turvey, 1999; Wann & Land, 2000; Wann & Swapp, 2000). Wann, Rushton, and Lee (1995) demonstrated that participants can steer accurately, using RF alone, if they are allowed to fixate their target. Li and Warren (2002) required participants to look at a point other than their steering target. They found that flow from the textured ground plane was not sufficient and that accurate steering required additional parallax information provided by vertical posts. This unusual gaze requirement negates the simple heuristics outlined in Wann and Land (2000) and forces the participant to recover something equivalent to instantaneous *heading*, which appears to be problematic using simple ground flow. In this article we address the relative roles of RF, ER information, and VD

---

[2] In the RR condition, for a fixed speed (V), target distance (Z), and trial duration (T), the heading angle (α) is related to the final angle between the fixation point and screen border (γ) by $\tan(\alpha) = \sin(\gamma)/(Z/VT-\cos(\gamma))$.

cues in steering curved paths toward a target, when participants are able to fixate their steering target.

Manipulating the Information Available for Heading and Steering Judgments

For heading and steering tasks we used a different, but equivalent, method of reproducing the conditions of Banks et al. (1996) using viewport motion (VPM) to manipulate the ER information. First, SR was added to the optic array using headings of up to ±12°. The rendered scene was then presented in a moveable viewport such that the whole display could be laterally translated to cause gain pursuit commensurate with the flow rotation present in the SR display or in a manner that conflicted with the display. The displays were presented in a dark environment, with no other spatial landmarks, so the ER component could be manipulated without changing the retinal pattern (assuming accurate pursuit). This method provided a means to reproduce ER signals of 100% and 50%, but also to introduce ER signals opposite to the RF rotation as a stringent test of the subtraction model (Royden et al., 1994). These conditions are shown in Figure 2, where manipulations that result in eye movements are labeled RR, and conditions that do not cause ER signals are labeled SR. The suffix of 1 or .5 indicates the 100% and 50% conditions, respectively. If the condition is created through the use of the viewport, then the label is suffixed with a *V*. As a consequence, a set of conditions (SR1, SR.5V, RR1V, NR.5V) kept the fixation target centered in the display so that the angle subtended across the ground plane to the left and right of the target remained equal and constant throughout the trial. This allowed us to assess the benefit of a VD cue across displays with equivalent RF patterns. We also used real-time gaze tracking to reduce the possibility of bias arising from gaze errors. In the follow-up experiments these methods were translated to a steering task. In the case of steering, VD information provides a strong target-centering cue. The VPM method allows control over the ER information while excluding the VD cue.

Li and Warren (2000) demonstrated that although heading errors may be high for an SR condition with a dot display, the errors were significantly reduced by additional parallax information. They suggested that the results with dot-flow fields have been due to the absence of sufficient parallax to resolve rotation components. We used ground textures with differing levels of geometric structure to explore whether the findings of Banks et al. (1996) could be accounted for with the Li and Warren explanation.

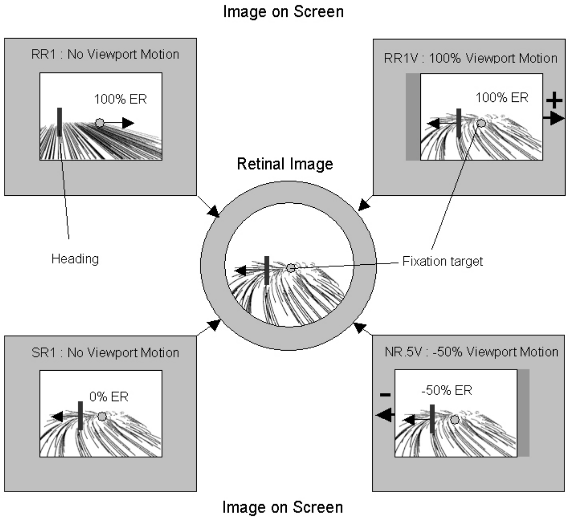

Figure 2. Top panels: Conventional real rotation display (RR1). If the observer tracks the target circle rotational flow is introduced into the retinal image and extraretinal (ER) gaze information is also available. Lower panels: If equivalent gaze rotation is simulated on the display and the target circle is fixated the same retinal image results. Moving the complete viewport manipulates gaze, to provide 0, 100% or –50% of the ER information that would normally be associated with the retinal flow pattern. SR = simulated rotation

Experiments and Predictions

All our experiments used realistically textured ground planes. White on black dot flow fields provide optimal contrast and under some conditions may even provide phosphor streaks indicative of flow. With simulations of natural ground surfaces there is an issue as to whether, given the spatial and temporal resolution of the display, a specific texture supports the perception of flow. We established three textures for use in our experiments[3]: an unstructured grass texture and a highly structured brick texture (with high-contrast linear elements and consistent parallel lines), both of which appeared to provide strong flow information and a faded texture that was designed to provide some impression of forward motion but made it difficult to detect flow vectors.

The first experiment involved a passive heading judgment task, where the participants travelled on a linear path but fixated an eccentric target. If the Banks et al. (1996) model holds, then heading judgments would be significantly more accurate when there is an ER correlate, even if this is introduced using VPM. We also expected inappropriate ER information (VPM in the wrong direction) to have a particularly detrimental effect. The experiment used the established brick and grass textures and two different locomotor speeds, 2.4 m/s (brisk walking speed) and 8 m/s (cycling or slow driving speed).

The second experiment investigated accuracy when actively steering toward a target at 8 m/s. If an ER model holds then, as proposed for Experiment 1, then VPM should help decode the rotated optic flow present in the SR condition to allow greater accuracy when steering. Quality of optic flow was manipulated using the established brick and grass textures. In principle, however, accurate steering could be performed without optic flow, using ER information alone or by using a visual reference for gaze motion. This was explored by using the faded texture. If ER information specifying gaze motion is the primary control variable, then a faded texture should not increase the steering errors.

The impact of RF was explored further in Experiment 3. We rotated the ground texture, during active locomotion, which introduced a bias into the optic/RF, without affecting ER signals. This provided a supplementary investigation as to whether active steering was influenced by RF, ER gaze information, VD information, or a combination of all three.

In Experiment 4 we added a more salient VD cue, similar to the emblem on the hood of a car, and introduced a subtle manipulation of this during trials. This tested the relative weighting that participants ascribed to RF or a strong VD cue and extends the debate of Rushton et al. (1998) and Warren et al. (2001) to the task of curvilinear steering.

General Method

All participants were unconnected with the study but had some general experience of viewing motion displays and making heading judgments. All had normal or corrected-to-normal vision.

The platform used was a PC with dual Celeron processors (Intel Corporation, Santa Clara, CA) running Windows 2000 and DirectX libraries (Microsoft Corporation, Redmond, WA). Images were presented in a light-excluded viewing booth with a large back projection screen (200 × 145 cm) providing a potential field of view (FoV) of 90° × 72°. A moveable viewport was created within this screen area. In Experiment 1 the viewport dimensions were 131 × 98 cm (66.5° × 52.2°) allowing 11.75° (23.5° total) of lateral VPM. The total lateral image–VPM across the screen varied from 5.24 cm (0.625°/s) for the smallest angle of heading to 23 cm (2.7°/s) for the largest. The dimensions of the viewport were reduced to 100 × 75 cm (45° × 36.9°) for the steering experiments (Exp. 2, 3, and 4) because the participants controlled the magnitude of lateral VPM, and greater latitude may have been needed. Images were rendered at a frame rate of 50 Hz and presented at a resolution of 1248 × 984 using an Electrohome 7500 graphics projector with fast phosphor tubes. Observers used both eyes to view the nonstereo image (bi-ocular viewing) at a distance of 1 m. Care was taken to ensure that invariant frames of reference were absent. The size of the screen placed the screen frame in the periphery (± 45°), but the screen edges were also covered in matt black tape, the room interior was matt black, and all incidental light was excluded from the projection

---

[3] In a heading judgment task with free gaze, the brick texture gave the best performance with relatively small errors (~3º). Grass with lower contrast, but visible flow, gave marginally worse errors than brick (~4º), but this difference was not significantly different, $t(5) = 1.37, p > .1$. The faded grass texture resulted in very poor heading judgments (~13º), indicating that although there was some flow information arising from this texture, it was not sufficient to support accurate perception of heading from a radial flow pattern.

booth. Between each trial a bright screen was flashed for 2 s to prevent dark adaptation by the observer.

The scenes presented comprised a ground plane tiled with a color image selected from a library of natural photographic textures (Figure 3) that provided an established amount of flow information.[4] Placed within the scene was a target (a single post for heading judgments or a pair of posts for steering), which was eccentric to the observer's initial heading and upon which the observer was instructed to fixate. The target moved in one of two ways: either as a result of observer's translation within the scene, which resulted in changes in both RF and gaze motion information, or by artificially moving the viewport in which the scene was projected, which only had a direct effect on gaze motion.

Head and body position were not stabilized, but an Applied Science Laboratories, (ASL; Bedford, MA) 5000 gaze monitoring system was used to check where the observer was directing gaze. The ASL system uses remote optics and pan-tilt tracking to follow the observer's head and eye and overlays point of gaze on the rendered scene, which could then be recorded on PAL SVHS tape, enabling the screening of trials for loss of fixation. In addition, we recorded the gaze co-ordinates for each rendered frame of the scene (50 Hz), which could be replayed at a higher resolution or frame-by-frame to check appropriate fixation.

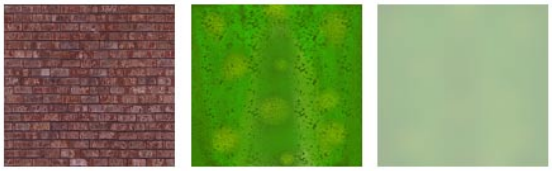

Figure 3. Examples of ground textures used in Experiments 1–3.

Experiment 1

The aim of this experiment was to reevaluate the findings of Banks et al. (1996) on the role of ER information in heading judgments. An additional factor was the use of VPM to create a wider range of conditions. In the experiment of Banks et al., the ends of the continuum were 100% RR and 100% SR. We recreate these under the labels of *RR1* and *SR1*, respectively:

> RR1: In this condition there was no SR component; hence, observer movement along an eccentric heading caused the target to move off to one side of the screen. When the observer tracked the target, the gaze of the observer rotated and thereby introduced a rotational component into the RF field. The task presented the observers with a curvilinear RF field, but they also had ER information available and therefore should make few errors in their final heading judgments.
>
> SR1: Here the scene camera was rotated prior to rendering each frame so that the target remained centered. The same flow pattern was presented to the eye as for RR1, but there was no actual gaze rotation and no ER information for the observer to use.
>
> RR1V: The addition of 100% VPM to an SR1 display meant that the whole image was displaced at each frame by an amount equivalent to the rotation that would have occurred in the pure RR1 condition, but without changing the flow components within the display. Hence, if observers tracked the target, their gaze would rotate at the correct rate for that specific heading condition (providing ER information). The difference from the RR1 condition is that the VPM method keeps the fixation target centered with respect to the display edges.
>
> SR1V: The SR condition also had a direct counterpart. We took the RR1 condition, where the target moved across the screen away from the direction of heading, but for each frame we moved the viewport back in the opposite direction an equivalent amount (RR − 100% VPM). This meant that the target remained stable and in front of the observer's bodyline, but the ground plane was displaced relative to the point of gaze and introduced a rotation into the flow field directly equivalent to SR1 (0% VPM). This also moved the display edges relative to the fixation points and reintroduced a VD cue to the SR condition.
>
> RR.5: Two additional intermediate conditions were matched: RR.5 recreated a condition of Banks et al. (1996) where the scene camera was rotated only half of the fixation angle and as a result an executed gaze rotation was required for the remaining 50% of the target motion. RR.5V produced the equivalent gaze requirement by adding 50%VPM to an SR1 display. The RR.5V

condition kept the fixation target centered in the display.

NR.5V: We introduced one final condition of SR–50% VPM, where the viewport was displaced in the wrong direction, thereby providing ER signals that conflicted with the direction of rotation in the flow field. We expected that this would make judgments very difficult.

The VPM method provided a test of varying amounts of ER information in the presence (RR1, RR.5, SR1V) or absence (RR1V, RR.5V, NR.5V, SR1) of the VD cue provided by the screen boundary.[5] The properties of each condition are summarized in Table 1.

Method

The display simulated observer translation across a textured ground plane, with a fixation point off the path. Two speeds of locomotion were chosen to simulate brisk walking (2.4 m/s equivalent to Banks et al., 1996) and steady cycling or slow driving (8 m/s, used in subsequent steering studies). At the beginning of each trial the fixation target was always centered in the display but was placed at either 12 m (for 2.4 m/s) or 24 m (for 8 m/s) to allow a range of rotation speeds and equivalent stopping distances. The eccentricity of locomotor heading was chosen to yield average gaze rotation speeds of 2, 4, 6, or 8°/s over a 2-s duration of travel (Table 2). To avoid participants becoming accustomed to the headings, an additional random jitter of ±0.5º was added to the heading for each trial.

The locomotion conditions (Table 2) were reproduced for each display condition (Table 1) and were repeated using both a brick and grass ground texture giving a total of 56 (4 × 7 × 2) conditions with 8 trials of each. To maintain consistent performance the display conditions (Table 1), speed of the locomotion conditions (2.4 or 8 m/s) and texture conditions (brick or grass) were varied in independent blocks, but the order of presentation was counterbalanced across participants. The two heading angles for the locomotion conditions were randomly interleaved. Eight participants, naive to the purpose of the experiment, completed the 448 trials.

Table 1

Display Conditions for Experiment 1.

Type of Rotn	Label	VPM condition	% Rotn within display	% Rotn in retinal flow* *(RF)*	% Extra-retinal rotn* *(ER)*	% Screen border ref *(VD)*
Real (Banks)	R1	RR +0% VPM	0	100	100	100
Real (VPM)	RR1V	SR +100% VPM	100	100	100	0
Simulated (Banks)	SR1	SR +0% VPM	100	100	0	0
Simulated (VPM)	SR1V	RR -100%VPM	0	100	0	100
50% Real (Banks)	RR.5	RR50% +0%VPM	50	100	50	50
50% Real (VPM)	RR.5V	SR +50%VPM	100	100	50	0
50% Negative	NR.5V	SR –50%VPM	100	100	-50	0

* assuming accurate target fixation & pursuit

Note. Rotn = rotation; VPM = viewport motion; ref = reference; VD = visual direction. See the text for an explanation of the experimental conditions in the *Label* column.

Table 2: Locomotion Conditions for Experiment 1

Speed (meters/second) : travel time (seconds)	Initial target distance (meters)	Heading (degrees)	Final target distance (metres)	Mean gaze rotation (degrees/second)	Min : max gaze rotation (degrees/second)
2.4 : 2	12	±6	7.2	2.0	1.2 : 3.3
2.4 : 2	12	±12	7.3	3.9	2.4 : 6.3
8.0 : 2	24	±6	8.1	6.0	2.0 : 16.9
8.0 : 2	24	±8	8.2	7.9	2.7 : 21.4

Note. Min = minimum; max = maximum

Observers were instructed to fixate and track a target for the duration of each trial. At the end of the movement phase, observers indicated where they perceived themselves to be heading using a joystick that controlled the lateral position of a vertical line within the visual scene. Sixteen practice trials of the control condition (RR1) were provided where participants received error feedback and became used to the display, controls, and eye tracker. Output of the ASL gaze tracker, overlaid on the render scene, was monitored during the experiment, and any trials where gaze did not track the target were repeated and participants were reminded of the fixation task.

To further ensure that the participants were tracking effectively, the gaze data were also recorded in synchrony with the locomotor data. The eye records were low-pass filtered at 6 Hz using Fourier transforms, and the prescribed and observed gaze behaviors were compared to compute a pursuit gain for each trial. For conditions that required target pursuit, trials were excluded from further analysis if the pursuit gain was less than 0.66 or greater than 1.2 or if there was evidence of regressive saccades. The ASL 5000 has an effective accuracy of ±0.5º for head-free tracking. With targets that have a sudden onset, large excursion, and exponential speed, this produces some quantization of the smooth pursuit records. The 0.66 and 1.2 criteria were arrived at after visually inspecting all 3,584 trial records; examples at the fringes of this window are displayed in Figure 4.

Results and Discussion

Gaze fixation. For the conditions that required static fixation (SR1, SR1V), trials were excluded from further analysis if the standard deviation of gaze position over the 2-s trial duration exceeded 1º. This resulted in exclusion of only 1.9% of the trials for SR1 and 2.1% of the trials for SR1V. For the remaining 98% of trials, the mean variation was 0.4º for SR1 and SR1V with the brick texture and 0.7º for SR1 and 0.6º with SR1V with the grass texture.

The pursuit gain criteria of 0.66 – 1.2 resulted in the following mean rejection rates: RR1, 4% ; RR1V, 5.5% ; RR.5, 8.6% ; RR.5V, 9.4% ; and NR.5V, 10.75%. For the remaining trials the mean gain for pursuit conditions was between 0.84 and 0.90, with no significant change across brick or grass textures (Figure 5). There was a significant variation across rotation speeds, however, with lower gains observed for higher rotation speeds: brick, $F(3, 21) = 49.2$, $p < .01$; grass, $F(3, 21) = 24.4$, $p < .01$. There was also a small but significant difference in the pursuit gains for the VPM conditions (RR1V, RR.5V: 0.89) as compared to their counterparts (RR1, RR.5: 0.85), $F(1, 7) = 12.78$, $p < .01$.

Heading errors. The mean difference between the heading indicated by participants and the actual heading was calculated in two ways. Constant errors were computed such that negative errors indicated an underestimation of the eccentricity of heading, irrespective of the direction of travel. Variable error was also computed as the within-subject standard deviation of heading judgments.

To address the issue of the ER contribution, the first comparison was across the conditions where there were differing proportions of gaze motion present, but the VD cue from the screen border was absent (RR1V, RR.5V, SR1, NR.5V, giving 100%, 50%, 0%, -50% ER). The upper panel of Figure 6 displays the signed errors across conditions for both brick and grass textures. It is clear that the reduction in ER information causes a systematic underestimation of the heading eccentricity, $F(3, 21) = 20.67, p < .01$, with the errors in the condition with gaze motion in the wrong direction (NR.5V) tending toward the wrong side of the fixation point. An interaction between condition and texture confirms that ground plane detail has little effect when the ER signals are strong, but when ER is removed or conflicting, the errors were greater for the sparser grass texture than for brick, $F(3, 21) = 6.69, p < .01$. There was also a significant effect of gaze rotation rate (Figure 6B), with errors increasing as the speed of pursuit increases, $F(3, 21) = 11.61, p < .01$. A notable feature of Figure 6B is that the errors that occurred in the 4°/s condition were equivalent to those of 8°/s and considerably greater than 6°/s. The 4°/s and 6°/s conditions are created with different locomotor speeds (2.4 m/s and 8.0 m/s), and this suggests that it is not the speed of rotation per se that creates a difficulty for the observer; rather, the rate of rotation relative to the translation component delineates heading. The analysis of variable error did not highlight any significant variation across conditions, texture, or rotation speed.

A second level of analysis was to compare the conditions where ER and VD cues were combined or selectively absent. Figure 7 presents the signed heading errors for the following conditions: RR1 (100% ER, 100% VD), RR1V (100% ER, 0 VD), SR1V (0 ER, 100% VD), and SR1 (0 ER, 0VD). When we coded the presence of ER and VD as factors in the analysis of variance, there was a significant effect of ER, $F(1, 7) = 29.59, p <.01$, and of VD, $F(1, 7) = 27.89, p < .01$, and a weak interaction, $F(1, 7) = 8.12, p < .05$. Once again there was no significant increase in variable error across the conditions. An equivalent analysis was performed using the 50% ER conditions (RR.5, RR.5V, SR1V, SR1), which produced equivalent results: $F(1, 7) = 8.69, p < .05$, for ER; $F(1, 7) = 18.82, p < .01$, for VD; and $F(1, 7) = 19.01, p < .01$, for their interaction.

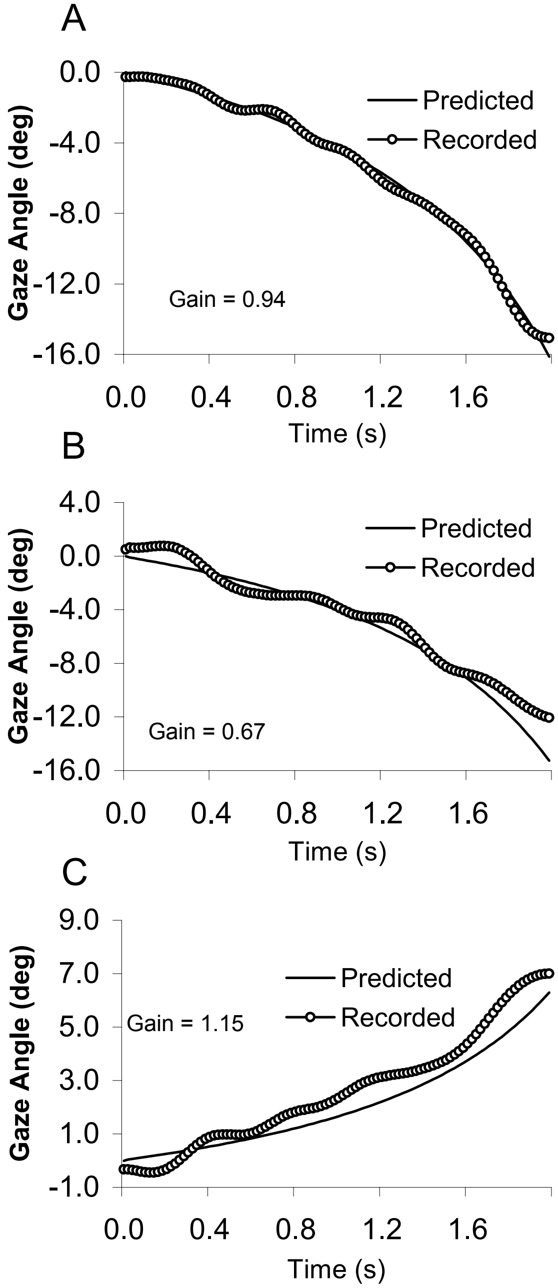

Figure 4. Examples of pursuit tracking by one participant. A: Ideal tracking performance. B: A small initial offset and some quantization of the recording can produce fluctuations that raise the gain above 1.1 or reduce it below 0.7. In both cases there is clear evidence of pursuit tracking, and the gaze position target remains with 1°. of the target.

These results suggest that both ER information VD cues assist heading judgments and that the removal of either results in an underestimation of perceived heading. The interaction highlights that when both ER and VD are removed (SR1), performance deteriorates beyond that predicted from a simple additive model.

Summary of Experiment 1. The general pattern of results reinforced the findings of Banks et al. (1996). The most accurate performance occurred in the conditions where the participants tracked a moving target and had veridical ER information. The results extend the findings of Banks et al., however, by demonstrating a statistically reliable pattern across a group of naive participants and demonstrating that the effects were not due to gaze pursuit errors. Our gaze pursuit results demonstrate that it is dangerous to assume that participants can perfectly track a moving fixation point. The condition with ER information in the wrong direction (NR.5V) was particularly useful in testing the additive (subtraction) model for ER input. It demonstrates that if an ER signal specifies gaze motion, but in the wrong direction, it is still factored into the heading estimation. This supports the view that something as unusual as the NR.5V condition is not interpreted as cue conflict and that performance is not just random; rather, it is systematically biased in the direction that would be predicted from a subtractive model (Royden et al., 1994).

When the ground texture had strong geometric features (brick: Figure 6A), errors for incompatible ER information were reduced as compared to an unstructured pattern (grass). These results suggest that highly structured flow that contains parallax information may assist in judging heading for cue conflict conditions where ER information is imperfect. This is consistent with Li and Warren (2000).

A further observation of interest is that it is not simply the rate of target rotation that made heading judgments difficult, but rather the rotation relative to the translation component. We found that participants had more difficulty with the experimental conditions of 4°/s at 2.4 m/s than 6°/s at 8 m/s. On reflection this is obvious: If the observer is creeping forward at 0.25 m/s, then a small amount of rotation masks the translation component, whereas an observer travelling at 20 m/s may tolerate considerable gaze rotation and still recover the linear components of the field. However, in previous studies locomotor speeds varied between 1.5 m/s and 4 m/s, and the results have been equated.

The second level of analysis compared the combined effect of the VD cue in addition to ER information. There was strong support for both ER and VD assisting in the judgment of heading, although when they were both removed (SR1) the errors were greater than would be expected for a simple additive model. This is also commensurate with Li and Warren's (2000) argument that SR1 is essentially a conflict situation, where ER and VD information suggest a different scenario to that presented in RF. The results displayed in Figure 7 suggest that adding either VD or ER can significantly reduce the error, and in part this may be due to conflict resolution.

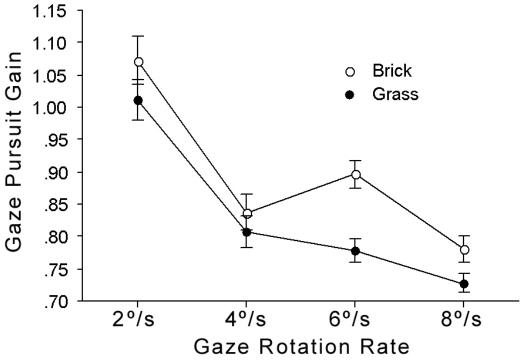

Figure 5. Gaze pursuit gain across speeds for brick and grass textures. Bars here and elsewhere (unless stated otherwise) indicate standard errors. See the text for an explanation of the abbreviations of the experimental conditions.

Experiment 2

Our aim in Experiment 2 was to translate the viewport paradigm used in Experiment 1 to the task of active locomotor steering to explore the use of RF, ER, and VD information. An RR condition with VD information (RR1, RF + ER + VD) was compared to one without (RR1V, RF + ER) and an SR condition that lacked both VD and ER information (SR1, RF only). Warren et al. (2001) proposed that the strength of the information arising from flow might affect the weighting attached to it. For this reason we used the three textures that had been previously tested to determine their strength.[6] Brick with strong linear geometry and grass with a random pattern enabled accurate linear heading judgments, whereas faded grass supported a percept of forward motion but not heading judgments.

In Experiment 1 we established that a stronger source of RF information (brick texture) resulted in more accurate heading judgments, even when ER information was degraded. The issue addressed with the faded texture was whether steering could be controlled accurately in the absence of salient flow information. In this respect the VD cue discussed in Experiment 1 also became more important. It has been proposed that locomotor steering could be completed purely on the basis of an estimate of egocentric VD (Rushton et al., 1998; Wann & Land, 2000). The VD of a fixated target relative to the bodyline could be detected from gaze position (ER), or in some settings there may be a retinal reference for this, such as vision of one's shoulders or handlebars. In the RR1 condition the truncated edge of the texture plane provided a visual reference equivalent to the edge of a car windshield. The distance of the target from the center of the display, defined by the display edges, or the ratio of the two angles from the target to the display edges, provides a direct indication of VD. In addition, Wann and Land (2000) pointed out that a curved path to a target could be achieved by maintaining a constant rate of closure for the target angle. This is also mirrored by a constant rate of expansion of the angle between the target and the display edge (or windshield). The VPM method used in RR1V removes the VD information by keeping the target centered in the display, equivalent to generating shutters that moved during each frame, such that equal proportions of the ground plane are visible to either side of the steering target (Royden et al., 1994; Wann et al., 2000). When combined with the faded texture this meant the RR1V condition provided ER information regarding target motion, but it did not have strong flow information or a visual reference to egocentric target direction.

Observers used a car steering wheel to actively steer to a target at 8 m/s (18 mph) using a simple point mass steering model for the VPM. This allowed the observers to pivot around their egocenter, similar to walking, but avoided the complications of learning the response characteristics of a specific vehicle. In a simulation such as this where there were no physical restrictions on turning arc or momentum considerations, participants could have completed a target alignment task in a single rapid pivot movement. This is seldom feasible in the natural world, where cars, bicycles, skates, and even running require a smooth change of trajectory. Hence participants were provided with practice trials to become accustomed to steering smooth trajectories. The task therefore required participants to make judgments as to whether their rate of turning was sufficient to bring them in line with the target some 2–6 s ahead.

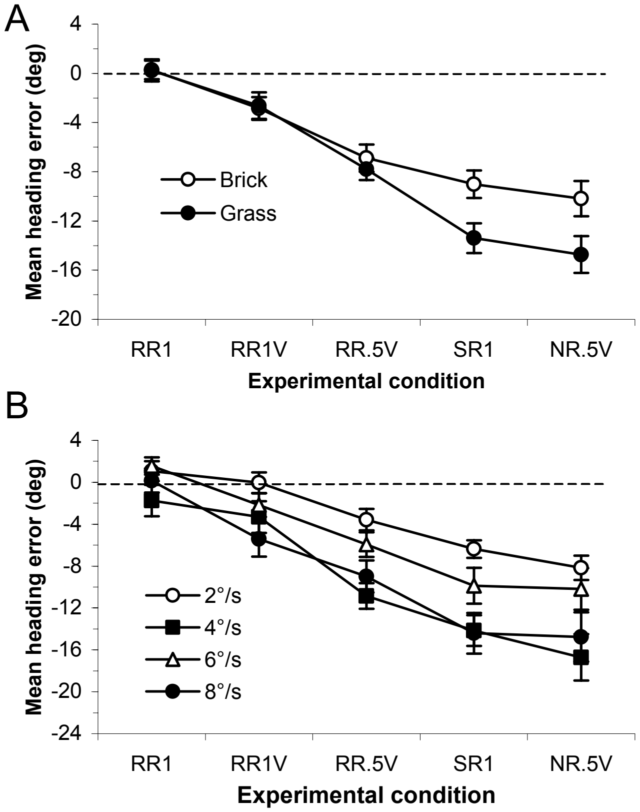

Figure 6. Signed heading errors (degrees) with different amounts of extraretinal (ER) information. A: Texture differences for brick (unfilled circles) and grass (filled circles). B: Different gaze rotation speeds. For both sets of data the first condition, RR1 (RF+VD+ER) is shown for comparison purposes and was not included in the statistical analyses. The other conditions are ordered along a continuum of reduced ER information: RRV (RF+100% ER), RR.5V (RF+50% ER), SR1 (RF+0% ER), and NR.5 (RF–50% ER). RF = retinal flow; VD = visual direction; ER = extraretinal. See the text for an explanation of the abbreviations of the experimental conditions.

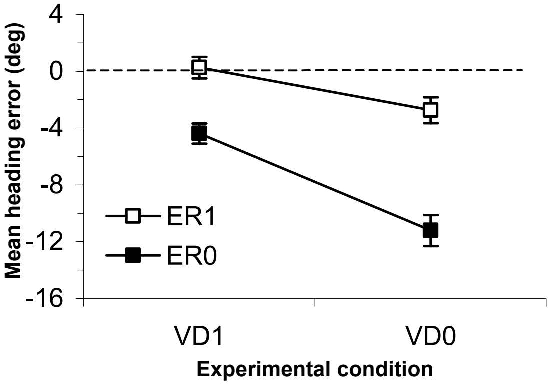

Figure 7. Signed heading errors (degrees) with different combinations of visual direction (VD) and extraretinal (ER) information: RR1 (RF+VD+ER), RRV (RF+ER), SRV (RF+VD), and SR1 (RF). RF = retinal flow. See the text for an explanation of the experimental conditions.

Method

The stimulus was a textured ground plane with a pair of white posts (2 m apart) offset to one side at an angle that varied between ±10°, ±14°, or ±18°. Three established textures were used: brick, grass, and faded grass. The 6 participants were instructed to steer a smooth trajectory that would take them through the posts. A self-centering force-feedback steering wheel was used, and practice trials were provided where an ideal, constant curvature trajectory was presented on the ground plane using the RR1 condition. The velocity was constant at 8 m/s. Trials started 60 m from the target and simulated locomotion until 10 m from the target. The trial was stopped before reaching the posts for three reasons. First, the target was sufficiently far away to negate binocular information as an issue. Second, it minimized the excessive visual rotation that occurs in the SR1 condition when nearing the target. Third, it prevented participants from receiving precise feedback as to the accuracy of their steering in each condition. We would also suggest that if one travels at 8 m/s and is not reasonably well-aligned with the gate posts by the 10-m point, then in the natural world only minor adjustments would be possible during the final second before arrival. The end steering error was calculated by fitting an arc of constant curvature for the 2 s immediately prior to stopping and extending it across the final 10 m to estimate the point that would be intercepted by the steering action 2 s prior to termination. The displays reproduced three conditions used in Experiment 1: RR1, RR1V, and SR1. The procedure for RR1V is technically more difficult here, in that the VPM at each frame has to be calculated on the basis of the current steering angle to induce the correct ER signals. Point of gaze was monitored using an ASL 5000 system as outlined in Experiment 1, but in this case instantaneous gaze motion is not prescribed and depends on the rate of steering. Pursuit gains were not calculated, but the output from the ASL 5000 was used to check compliance with the fixation conditions. Six naive observers participated.

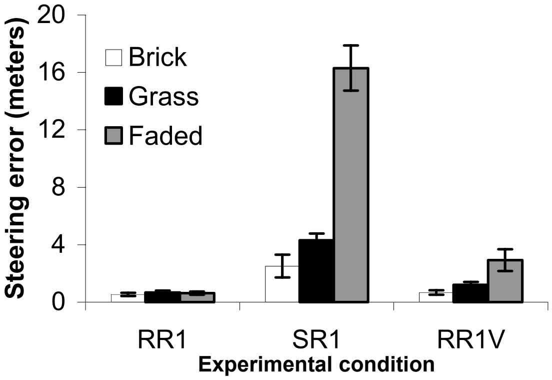

Figure 8. Mean steering errors (meters) for Experiment 2 with real gaze rotation (RR1) or simulated rotation with viewport motion (VPM; RRV) or without VPM (SR1). Bars indicate standard errors.

Results: The Adequacy of Flow

Steering errors were calculated as the lateral offset from the center of the gate in meters and are displayed in Figure 8. The data show that there were consistently small steering errors when all information was available (RR1), and the mean error of 0.6 m (*SE* = .12) would have resulted in safe passage through the gate (2 m wide). This also serves to confirm that the error estimation procedure was appropriate with calculated errors of ~3º. Accurate steering was seen for RR1 irrespective of the texture placed on the ground plane, *F*(2, 10) = .609, *ns*. The faded grass texture was previously identified as being insufficient for supporting even simple judgments of heading from flow. In the RR1 condition, however, the accuracy with the faded texture was equivalent to that of the other textures. This suggests that participants could steer accurately without flow, by using ER or VD information.

In the RR1V condition the participants had access to appropriate ER information, but unlike the RR1 condition the target gate remained centered within the viewport and did not move toward the edge of the texture plane. This removed the VD cue provided by the screen boundary. The advantage of brick and grass over faded grass for RR1V suggests that, in the absence of VD information, RF does assist in judging steering.

The errors were consistently large when both ER and VD information were removed (SR1). As expected, errors were especially elevated for the faded grass texture where there is very little directional information. The brick texture did display an advantage over grass texture for SR1, $t(5) = 2.156$, $p < .05$, and this may indicate an advantage of its linear geometry in perceiving rate of rotation.

Figure 9 displays sample trajectories for one participant to illustrate the effect of removing the VD, ER, and degrading the flow information.

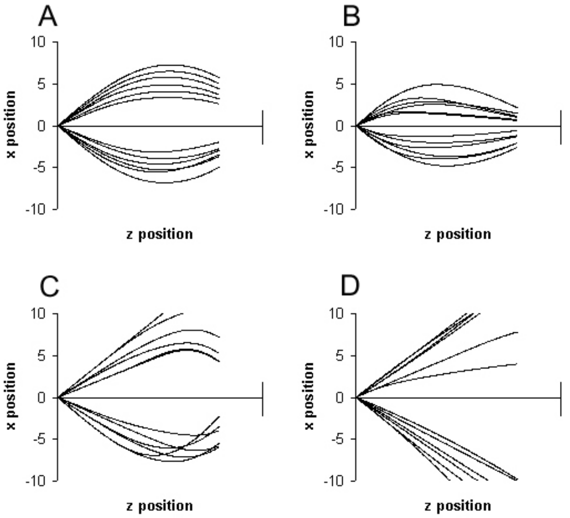

Figure 9. The course of 12 paths of different initial angles for a single participant. Trajectories start at the left and terminate 10 m from the target indicated by the vertical line. Contrasting steering behavior is visible as information is removed or degraded. A: RR1: extraretinal (ER), visual direction (VD), and brick retinal flow (RF). Accurate, smooth, rounded paths are followed when strong RF, VD, and ER information is available. B: RR1V: ER and brick RF. When VD information is removed, steering is no less accurate, but the paths followed are more direct, indicating coarser control over navigation. C: SR1: Brick RF. When both VD and ER information is removed, steering is much less accurate. The paths are straight during the first half of the trials indicating that the participant is not receiving sufficient information from RF to determine the necessary steering requirement. Eventually the rotation in RF (caused by simulated fixation of the target) becomes unmistakable, and the participant's attempts to compensate cause erroneous understeering or oversteering. D: SR1: Faded RF. The degraded RF does not supply sufficient information to detect the increase of the rotation component within RF, and so steering behavior is negligible under these conditions.

Discussion: The Role of Flow

Experiment 2 demonstrated that RF on its own, without appropriate ER information, was not sufficient for accurate steering around a curved path (SR1 condition). In the RR1 condition two sources of information are present in addition to flow: The displacement of the target gate relative to the observer induced visual pursuit and therefore ER gaze information. The same displacement can also result in a visual (retinal) referent for rate of change of VD, by moving toward the edge of the windshield or in this case the boundary of the ground plane. When these information sources were available, steering was shown to be highly accurate even when RF was extremely degraded. It has been proposed that the rate of change of VD, on its own, is sufficient to guide steering (Wann & Land, 2000). When the latter cue was removed (RR1V), errors increased.

In summary, the results support an argument that ER information is important for accurate steering. However, there is still evidence for the role of flow: The errors in RR1V were greater in the faded condition than for the other textures and there was some indication that different texture types (brick vs. grass) may influence performance. The role of flow and ER information, however, may be reduced if a visual referent for target direction or motion is available. These results are in agreement with the findings of Experiment 1 and support a model of steering control where RF, ER, and VD information is selectively weighted ($\beta_{1\text{–}3}$) depending on their availability and salience.

This simple expression will form the basis for a full steering model developed in the Appendix:

$$\text{Steering requirement} = f\,(\beta_1 RF,\ \beta_2 ER,\ \beta_3 V \quad (1)$$

Experiment 3

The purpose of Experiment 3 was to further explore the combinatory model proposed in Equation 1. The results in Figure 8 provided clear support for the roles of VD and ER information. The results from the RR1 task with the faded texture suggest that the flow information may be redundant if VD and ER are available to guide steering. If RF is being used, then a perturbation of the flow field should produce predictable steering errors. To introduce a subtle bias into the flow field, we rotated the ground plane around the point of gaze (the steering target) during locomotion. Although the origin of gaze rotation is the observer's head, ground fixation dictates that the rate of rotation is adjusted to keep the fixation target stationary within the flow field. This is geometrically equivalent to the observer's eye rotating around the point of fixation at an ever-increasing rate, as if attached by a cable of decreasing length (Figure 10). This gives rise to an asymmetrical, expanding spiral motion pattern in the retinal image. Adding a rotation of the ground plane around the point of fixation, therefore, makes the spiral in the retinal image tighter, equivalent to what might occur if the participant was travelling on a more eccentric path. The manipulation still satisfies the constraint that the texture elements remain stationary at the point of fixation and the steering target still appears to be fixed to the ground plane. The ER and VD information is unaffected by the rotation, and hence there is a dissociation between the three sources of information. The manipulation differs from a virtual prism (Warren et al., 2001) in that a simple prism would not dissociate the velocity components of ER and VD from RF.[4] A simulation of these conditions is presented for the model in the Appendix.

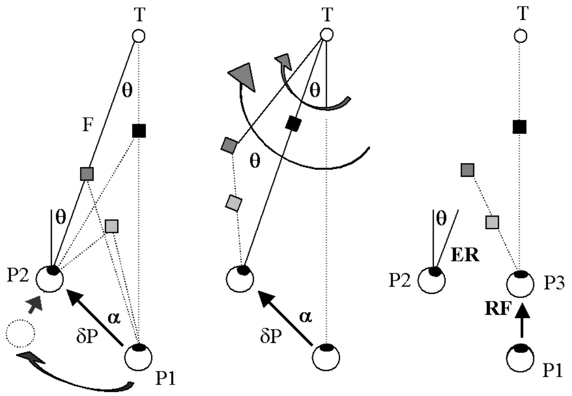

Figure 10. Dissociating retinal and extraretinal (ER) sources of locomotor information. Left: If an observer moves by δP from Point (P) 1 to P2 while fixating target T, this is geometrically equivalent to the observer undergoing a pure rotation around T with a coincident translation along the radius of rotation. The angle of gaze rotation θ is dictated by the distance travelled δP, the heading angle α, and resultant fixation distance F. Position P2 and gaze angle θ result in a change in the projection angles for ground objects (squares) giving rise to curvilinear retinal flow (RF). Center: If the ground plane is rotated around T by angle θ during the translation, then this can null the rotational component of RF. Right: The RF experienced in the center diagram would be equivalent to motion directly toward target T, from P1 to P3 ($\alpha = 0$), but the eye has still moved to P2 so the ER information is still specified by θ.

Method

The steering conditions were identical to those used in Experiment 2. The grass texture was used and the RR1 (real gaze rotation) and RR1V conditions were repeated. Both of these conditions present strong RF and ER information, but the RR1 condition includes a VD cue that is not available in the RR1V

---

[4] Contrary to what is sometimes assumed, a prism does not displace the FoE of flow from the target. If the observer fixates the target, then the FoE is still centered at fixation, but the flow field is weakly curved. The curved locomotor trajectory of Rushton et al. (1998) arises because of a centering strategy for when participants try to walk a straight path. For trajectories that are intentionally curved, the rate of change of VD and ER is unaffected by a virtual prism, and the only bias introduced may be in the final adjustments very close to the target.

condition (see the *Discussion* section for Experiment 2). The ground plane was rotated around the target object (the center of a pair of posts 2 m apart). The rate of rotation was a function of the distance of the observer from the target and used a polar estimate of the gaze rotation ($\dot{\theta}$) that would result from a heading error of α (Figure 10):

$$\dot{\theta} = V \sin(\alpha)/F, \tag{2}$$

where V was the locomotor speed (8 m/s), F was the line of sight distance calculated every frame, and α was chosen to be 20 º. Rotation away from the path (labeled *positive*) and rotation toward the path (*negative*) were counterbalanced with trials where there was no ground rotation. One caveat is that degree of bias introduced into RF depends on the participants' position and rate of steering; it was not possible to predict the precise steering error, but the direction of the error should always be opposite to the direction of induced ground rotation. Trial duration, initial headings, speed, and steering device were identical to Experiment 2. Steering errors were calculated using the procedure in Experiment 2: A constant curvature arc was fitted to the first third (2 s) of the trajectory and extrapolated to estimate where each participant would end if he or she maintained that rate of steering. The procedure was then repeated for the next third (2 s) and the final third (2 s), to provide 3 error estimates for the early, mid, and late steering response. Six naive observers completed the experiment.

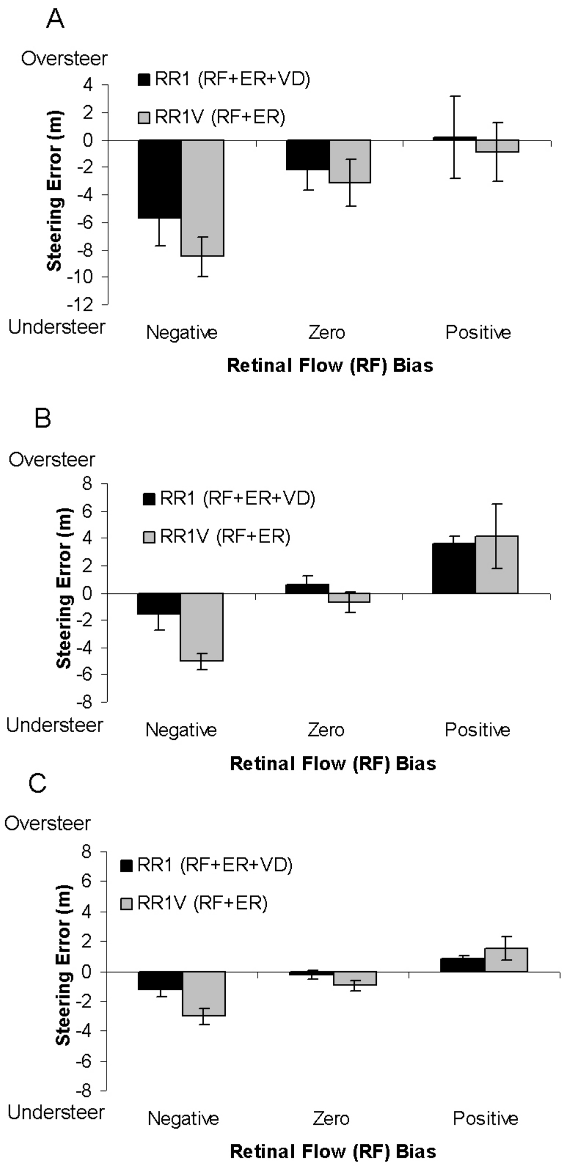

Figure 11. Mean steering errors for Experiment 3 with real gaze rotation with (RR1) or without (RR1V) visual direction (VD) information. A: The first 2 s (60 – 43.3 m) of the trajectory. B: The middle 2 s (43.3 – 26.7 m) of the trajectory. C: The final 2 s (26.7 – 10 m) of the trajectory. See the Experiment 1 section of the text for an explanation of experimental condition abbreviations. Bars indicate standard errors. ER = extraretinal.

Results and Discussion

Observers using RF while negotiating a bend should be steering to reduce the rotation in the flow field. If the ground plane is rotated away from their path, they should perceive that they are not turning at a sufficient rate and increase their steering angle (oversteer). If the ground plane is rotated toward their path, they should perceive that they are turning too quickly and decrease their steering angle (understeer). Figure 11 displays the errors for the early (A), mid (B), and late (C) steering response that resulted when the ground plane was rotated either toward the path (negative) or away from the path (positive), along with trials where there was no rotation. During the first 2 s there was a tendency toward understeering, but the direction of rotation significantly biased the response in the predicted direction, $F(2, 10) = 14.39$, $p < .01$. There was no difference between the results of the RR1 and RR1V conditions. By the middle 2 s the steering response was well judged with a mean error of under 1 m for the stable ground conditions, but a bias in the predicted directions for the ground rotation conditions, $F(2, 10) = 13.48$, $p < .01$. The pattern was the same for the final 2 s, $F(2, 10) = 16.91$, $p < .01$, except there were higher errors in the RR1V condition as compared to the RR condition, $F(1, 10) = 9.31$, $p < .03$. Coding the rotation factor with coefficients of orthogonal polynomials to assess understeer and oversteer confirmed that for all three segments the effect was linear ($p < .01$) with no quadratic component. These results clearly demonstrate that ground flow is being used during early, mid, and late steering responses. Other cues were unaffected by the ground flow manipulation and cannot explain the strong bias toward understeering and oversteering that resulted. The difference in the RR and RR1V conditions during the final steering segment is of note. These two conditions are equivalent except that the target stays centered in the viewport for the RR1V condition and this removes the VD information. The lack of difference in conditions for the first 4 s suggests that steering was based predominantly on flow + ER information at this time. During the final 2 s, however, the target poles would move more rapidly toward the screen edge (condition RR1) in response to any steering error, which suggests that VD may become salient only as the observer approaches the target.

The errors observed in Figure 11C can also be compared to those predicted by the steering model in the Appendix for the same conditions. The model's parameters of μ, υ were set on the basis of the control condition, and the weightings (β_1, β_2, β_3) were assumed to be equal. When the biased inputs were simulated, the model steering errors lay between 1.8 m and 2.2 m for the RR1 condition and 3.0 m and 3.6 m for the RR1V condition, which are equivalent to the empirical results of Figure 11C.

Experiment 4

In the final manipulation, we sought to establish the relative weighting attached to RF and VD information during curvilinear steering task. Rushton et al. (1998) proposed that for walking a linear path, VD is the primary cue, whereas Wood et al. (2000) and Warren et al. (2001) have proposed that there is a combinatory system. We revisited this issue with a curved steering task taken from Experiment 3. The display condition was identical to condition RR1V (RF + ER) except that an explicit VD reference similar to the emblem on the hood of a car was provided (Figure 12). By making the emblem drift to the left or right (toward or away from the target destination), positive and negative VD bias could be introduced. At the same time we introduced RF bias using the method from Experiment 3 (Figure 10). The RF and VD stimuli could be independently manipulated to promote understeer or oversteer, and the steering behavior that resulted demonstrates the manner in which these two sources of information interact.

Method

As in Experiment 3, texture rotation was used to add bias to the RF. A high-contrast gravel texture, halfway between the grass and the brick, was mapped onto a ground plane and was used to simulate driving over a road. This was then rotated around the target posts either toward or away from the required direction of steering using the same parameters as Experiment 3 (Equation 2). A badge presented within the scene provided an explicit VD reference for the center of the locomotor vehicle. The veracity of this reference was manipulated by making it drift to the left or right at a rate of 2.4º/s (15º in total). The rotation of the ground plane and rotation of the badge could be in equivalent or opposite directions. The rotational conditions were labeled on the basis of whether they would result in oversteer (positive) or understeer (negative) if the response was based purely on that specific input. This yielded trials with three levels of texture (RF) rotation (positive, zero, negative) and three levels of the car emblem (VD) rotation (positive, zero, negative). There were therefore 66 trials in total, made up of 12 control trials (with no bias) and 54 randomized trials. Two initial heading angles were used of ±10º and ±14º with an additional jitter of ±1º.

Figure 12 shows an example of the display used, with textured ground plane, fixation target (two vertical posts), and badge. Participants were instructed to keep their gaze fixed on the posts and then steer a smooth course that would pass between them. They pressed a button on the steering wheel to initiate each trial. They were told that some trials might make them feel as if they were sliding or skidding, but not to be concerned by this and to attempt to maintain a smooth and accurate course.

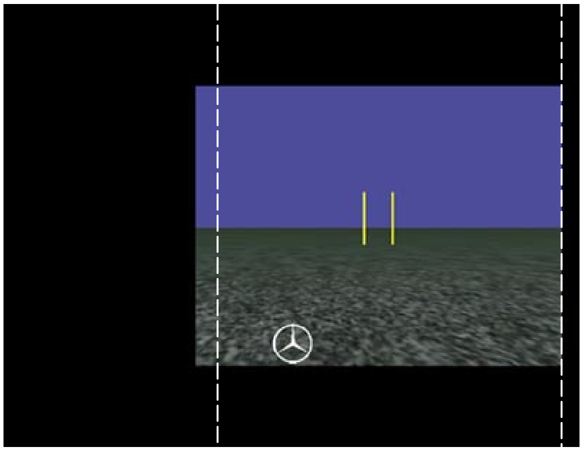

Figure 12. The display had black shutters that maintained the target posts centrally within the viewport (highlighted in the diagram by the dotted lines), which removed the visual frame of reference that the static edge would otherwise have presented. A white Mercedes badge was then used to reintroduce this visual direction information explicitly. The badge was then moved to one side or the other during some conditions to create biased drift.

Results and Discussion

The mean steering errors displayed in Figure 13 show a pattern of results that are consistent with both RF and VD information influencing steering behavior.

When veridical RF was available (central group of 3 bars), bias introduced into VD resulted in a directional error commensurate with the use of VD information, $F(2, 5) = 8.58$, $p < .01$. The equivalent comparison is to consider the errors when the VD information is veridical but RF is biased (center bar of each of the 3 groups). It is clear that bias introduced into RF resulted in a directional error commensurate with the use of RF, $F(2, 5) = 74.27$, $p < .01$. Because biased RF in the presence of veridical VD and biased VD in the presence of veridical RF both cause erroneous steering, we can infer that the system does not prioritize one particular source of information. We can also examine the conditions where bias is present in both VD and RF simultaneously (Figure 13, flanking sets of bars). In these situations the resultant steering errors showed a regular and systematic pattern. When the bias in both sources of information occurred in the same direction, the magnitude of steering errors increased, while conflicting bias reduced steering errors. For example, positive VD bias increased the size of steering errors resulting from positive RF bias but decreased the degree of error resulting from negative RF bias. This pattern is equal and opposite for negative VD bias; it accentuated errors for negative RF bias and reduced errors for positive RF bias.

The results support a simple additive model for the steering system, where RF and VD information carry a comparable weighting. Across the conditions there was an error gradient of 0.2 m/degree RF bias and 0.17 m/degree of VD bias. Neither source of information overrides the other in this competition. It seems that even with a VD cue as explicit as a car emblem, the flow information was still strong enough to alter steering behavior. We have not been able to independently manipulate ER cues within this experiment, but the ER gaze angle information provides a reference equivalent to the veridical VD condition across all conditions. However, it may be the case that bias in the VD information alters the perception of gaze angle, so the influence of ER cues cannot be separated at this stage.

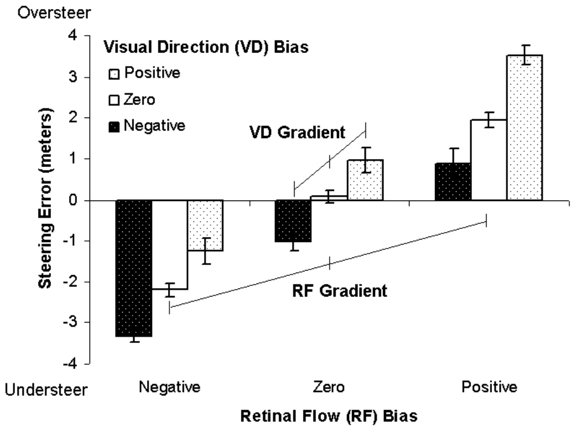

Figure 13. Mean steering errors (meters) from 6 participants showing the effect of contrasting combinations of RF and VD bias for the final 2 s of the trajectory.

General Discussion

The aim of these experiments was to revisit the paradigm of Banks et al. (1996) and translate it to the more ecological task of locomotor steering. Experiment 1 used a moving viewport paradigm to re-examine the claim that ER information is necessary to enable accurate judgments of locomotor heading. The position of Banks et al. was supported, but we provided stronger evidence for a combinatory model, through the introduction of negatively correlated ER information (condition NR.5V). The moving viewport technique also allowed manipulation of a frame cue to VD. Although this cue is not in itself sufficient to specify heading, it significantly improved the judgment of heading from flow, in the absence of ER cues. It might be considered that the VD cue was an artifact of the visual projection, so some comment on its ecological validity is warranted. In everyday tasks, whether locomotor or static, we have a peripheral view of our body and limbs. In a car a strong visual reference is provided by the edges of the windshield, and in a virtual environment there are the screen edges of the head-mounted display or projection volume. Only in limited settings, such as steering to a target in the dark, is this information removed. We propose that this type of information is present in many locomotor settings and should be considered as a potential contributor to locomotor control as specified in Equation 1.

Experiment 2 translated the viewport paradigm to the task of steering a curved path, across open texture, toward a gateway. We found that a pure RF field, as presented by the SR condition, was not sufficient for accuracy steering. This seems in conflict with the findings of Wann et al. (1995), but in this earlier study participants could use any steering strategy and the task could be simplified by crossing the target line, repeatedly adjusting their direction of travel. In the current experiment they were required to execute a smooth curved path (Figure 9A), and RF on its own lead to significantly greater errors, missing the target posts by more than 3 m. When VPM was added to the SR display this displaced the target gate to the side, depending on the steering action, but the target gate remained centered in the viewport. This introduced a pure ER signal (with no VD cue) and this significantly improved steering performance. The best performance was achieved in the RR1 condition where RF, ER and VD information was available. Surprisingly, accurate performance (0.5 m lateral error in steering to a 2 m gate) was maintained even when RF was severely degraded (faded texture), which suggests that ER and VD information may be sufficient on their own. When VD information was removed with the faded texture, errors increased significantly. This could be interpreted as evidence for a pure VD strategy (Harris & Rogers, 1999; Rushton et al., 1998), but we would council against this interpretation. The steering errors in RR1V (no VD) were equivalent to RR1 (with VD) when there was a suitable texture on the ground, which suggests that VD information can be removed if there is strong RF that can be combined with ER information.

A central role of RF was confirmed in Experiment 3, where the ground plane was rotated around the point of fixation to introduce an RF bias, without altering the ER or VD information. The result was a significant, reciprocal bias toward understeer or oversteer throughout the trajectory depending on the direction of texture rotation. Experiment 4 then introduced an equivalent bias into VD information to investigate the relative contribution of RF and VD information in specifying understeer or oversteer. The results of these steering experiments strongly support the role of RF, but they also demonstrate the function of VD and ER information when controlling of steering. The role of ER information is not just to "correct" RF because in principle ER information could be viewed as a nonvisual estimate of rate of target closure (Wann & Land, 2000). The results of Experiment 2 indicated that when the only salient information was ER (RR1V, faded texture), steering performance was still accurate (1 m error).

The results of Experiments 3 and 4 provide a clear confirmation of the use of RF during steering, in the presence of other retinal (VD) or ER sources of information. How RF is used is not yet clear, and we would warn against the assumption that *heading* is the control solution. We have demonstrated that RF can specify whether a trajectory takes the observer to a fixated target without recourse to a percept of instantaneous linear heading. In this analysis if the observer fixates a target (as in these experiments), an appropriate curved path results in straight (but nonradial) flow lines, whereas understeer or oversteer curves the flow lines reciprocally (Kim & Turvey, 1999; Wann & Land, 2000; Wann & Swapp, 2000). The texture rotation we introduced in Experiments 3 and 4 would have caused flow curvature of a type that would indicate a steering error to an observer using this model.

The challenge for future research is to establish how redundant information such as RF, ER, and VD information is combined and the weighting attached in different settings. We make an initial proposal in the Appendix, by combining them in a second-order model, based on the proposals of Fajen and Warren (2003), that acts as a point attractor toward the point of fixation. If the observer fixates the intended target, RF, ER, and VD information can all be used to provide independent estimates of understeer or oversteer. As illustrated in the Appendix, the perceptual inputs do not require precise estimates of rotational speed ($< 4°/s$) and can even be reduced to estimates on an interval scale. Under this scheme each input from RF, ER, and VD could be classified as indicating oversteer (1, 2, 3) or understeer (–1, –2, –3), thereby simplifying the combinatory process. This is equivalent to distinguishing between a minor steering error of a few degrees and a major steering error that would take us off the road. There is the latitude to make parameter estimates more graduated, but the Appendix illustrates that the problem can be reduced to a simple combination rule that should be robust when multiple sources of information are available. If fixation of the steering goal is maintained, then RF, ER, and VD are all correlates of $V\sin(\alpha)/F$ (Figure 9), so the slopes of the function are shallow when F is large and steepen as the observer approaches the target or veers wildly off course. As a result, the expanded version of Equation A1 specifies a surface where the information channels the observer along the ideal (zero state) trajectory.

One limitation of this model is that it only acts as a point attractor to the object that is fixated by gaze; it does not handle the case of fixating a point other than the target (Li & Warren, 2002). A second limitation is that the model is not responsive when the fixation point is very far away, when there is only a very small rotation of viewpoint. Whether this is completely consistent with human performance in different natural steering tasks is an empirical issue that is not yet resolved.

Harris and Rogers (1999) stated: "We challenge flow researchers to provide some compelling evidence for a significant role of optic flow in the control of direction of locomotion on foot" (p. 449). Warren et al.(2001) would claim to have answered that "on foot," but we have reinforced this in a task more akin to steering a vehicle. The task we have presented could be considered as simple "rotation nulling," because to steer accurately observers must bring the target gates in front of them. We argue that the steering mechanism was based on a systematic reduction of the combined rotation components obtained from RF, ER, and VD information. On the simple application of this approach there is no requirement to reconstruct the three-dimensional structure or a planned path within the three-dimensional environment. This does not remove the ability to plan a path. If there are a series of waypoints to be navigated, then the planning component is incorporated into how the driver fixates each of these locations. A simple example is seeking out the tangent point for a smoothly curving road (Land & Lee, 1994), but this principle can also be extended to more general trajectories of varying curvature: If a driver wishes to take a "racing line" and follow a path through the apex of a corner, then in our model the skill would be in learning to fixate the correct via points that act as knots to spline together the trajectory segments (Wann & Land, 2000). Hence, planning a course, or avoiding an obstacle, may be accomplished at the level of where you look and when, but active steering remains under the control of a simple point attractor that "nulls rotation." It is also appealing that this model concurs with advice from experienced road users, that to negotiate and obstacle you should look to the side that you wish to pass, because if you look at an obstacle you tend to hit it (Experienced Motorcycle Rider Course, 1992). We do not deny that these tasks could be completed by recovering *heading* and its rate of change, but this would still leave the problem of how the observer uses a heading angle to plan a path in three-dimensional space. Fajen and Warren (2003) proposed an equivalent point attractor model that acts to null the heading offset, but this has the additional requirement of recovering target distance. Our approach does not require the estimates of instantaneous heading or target distance, and in this respect we consider it a parsimonious model for "simple" steering that has the flexibility to be extended to complex, planned trajectories through the use of active gaze.

References

Banks, M. S., Ehrlich, S. M., Backus, B. T., & Crowell, J. A. (1996). Estimating heading during real and simulated eye movements. *Vision Research*, *36*, 431–443.

Crowell, J. A., Banks, M. S., Shenoy, K. V., & Andersen, R. A. (1998). Visual self-motion perception during head turns. *Nature Neuroscience*, *1*, 732–737.

Ehrlich, S.M., D.M. Beck, J.A. Crowell, T.C.A. Freeman, and M.S. Banks, *Depth information and perceived self-motion during simulated gaze rotations.* Vision Research, 1998. **38**(20): p. 3129-3145.

Fajen, B. R. (2001). Steering toward a goal by equalizing taus. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 953–968.

Fajen, B. R., & Warren, W. H. (2000). Go with the flow. *Trends in Cognitive Sciences*, *4*, 319–324.

Fajen, B. R., & Warren, W. H. (2003). Behavioral dynamics of steering, obstacle avoidance, and route selection. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 953-968

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, *49*, 182–194.

Harris, J. M. (2001). The future of flow? *Trends in Cognitive Sciences*, *5*, 7–8.

Harris, J. M., & Rogers, B. J. (1999). Going against the flow. *Trends in Cognitive Sciences*, *3*, 449–450.

Kim, N. G., & Turvey, M. T. (1999). Eye movements and a rule for perceiving direction of heading. *Ecological Psychology*, *11*, 233–248.

Land, M.F. and D.N. Lee, *Where we look when we steer.* Nature, 1994. **369**: p. 742-744.

Lappe, M., Bremmer, F., & van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, *3*, 329–450.

Lee, D. N., & Lishman, R. (1977). Visual control of locomotion. *Scandinavian Journal of Psychology*, *18*, 224–230.

Li, L., & Warren, W. H. (2000). Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. *Vision Research*, *40*, 3873–3894.

Li, L., & Warren, W. H. (2002). Retinal flow is sufficient for steering during observer rotation. *Psychological Science*, *13*, 485-491

Motorcycle Safety Foundation, Inc. (1992). *Motorcycling Excellence: Skills, Knowledge, and Strategies for Riding Right* (H. Cadermatori, Illus.). North Conway, NH: Whitehorse Press.

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

Royden, C. S., Crowell, J. A., & Banks, M. S. (1994). Estimating heading during eye movements. *Vision Research*, *34*, 3197–3214.

Rushton, S. K., Harris, J. M., Lloyd, M. R., & Wann, J. P. (1998). Guidance of locomotion on foot uses perceived target location rather than optic flow. *Current Biology*, *8*, 1191–1194.

Rushton, S. K., & Salvucci, D. D. (2001). An egocentric account of the visual guidance of locomotion. *Trends in Cognitive Sciences*, *5*, 6–7.

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

van den Berg, A. V. (1993). Perception of heading. *Nature*, *365*, 497–498.

Wann, J. P., & Land, M. (2000). Steering with or without the flow: Is the retrieval of heading necessary? *Trends in Cognitive Sciences*, *4*, 319–324.

Wann, J. P., & Land, M. (2001). Heading in the wrong direction. *Trends in Cognitive Sciences*, *5*, 8–9.

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. G. Bardy, R. J. Bootsma & Y. Guiard (Eds.), *Studies in perception and action III* (pp. 171–174). Mahwah, NJ: Erlbaum.

Wann, J. P., & Swapp, D. K. (2000). Why you should look where you are going. *Nature Neuroscience*, *3*, 647–648.

Wann, J. P., Swapp, D., & Rushton, S. K. (2000). Heading perception and the allocation of attention. *Vision Research*, *40*, 2533–2543.

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

Warren, W. H., & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America*, *7*, 160–169.

Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience*, *4*, 213–216.

Warren, W. H., Mestre, D. R., Blackwell, A. W., & Morris, M. W. (1991). Perception of circular heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 28–43.

Wood, R. M., Harvey, M. A., Young, C. E., Beedie, A., & Wilson, T. (2000). Weighting to go with the flow? *Current Biology*, *10*, R545-R546.

(Appendix Follows)

Appendix

A Steering Model Based on Redundant Visual Inputs

If an observer fixates a steering target that is eccentric to the current locomotor path, then a gaze rotation occurs. The rate of gaze rotation is a function of the locomotor speed (V), the eccentricity (α) of the target, and distance (F) of the target from the locomotor path. The gaze rotation may be detected through ER signals; if there is a visual reference for the locomotor axis, it may be detected as a changing VD. The gaze rotation is also reflected in a reciprocal rotation of ground features in RF, around the point of fixation, which is stationary on the fovea (Figure 10; see the text). There are, therefore, three potential sources of information through which gaze rotation can be detected. These can be combined using a simple point attractor (Fajen & Warren, 2003), to provide a robust model for visually guided steering:

$$\ddot{\theta} = \mu(\beta_1 RF, \beta_2 ER, \beta_3 VD) + \upsilon\dot{\theta}. \quad \text{(A1)}$$

In this equation, $\dot{\theta}$ specifies the angular velocity and $\ddot{\theta}$ the angular acceleration of the observer (rads/second and rads/second/second, respectively); μ is a constant that dictates the response speed and υ is a damping factor, which may in some cases be dictated by the physical properties of the locomotor device (e.g., steering rake). The balance of μ and υ dictate the stability of the system and can be set for the initial task conditions. Values of μ and υ must be selected that close down the initial angle in the distance or time available, and υ must be sufficient to avoid oscillation of the system. The perceptual inputs of RF, ER, and VD (rads/second) are estimates of the degree of rotation of the target, or ground, relative to the observer weighted by an individual β component that here is assumed to be *1/number of components* ($\beta_1 = \beta_2 = \beta_3 = {}^1/_3$). Equation A1 specifies the steering action required through any one of three sources of visual information, or the combination of all three. High precision estimates of RF, ER, and VD are not essential. Figure A1 uses the task conditions of experiment

3 to demonstrate the response of the model when the sensitivity is degraded from 1°/s to 4°/s. With less sensitive settings the model travels a further distance before responding and then steers appropriately and closes down the target angle. The average displacement errors for the settings of 1°/s and 4°/s are 0.32 m and 0.48 m. It is interesting that for both settings the steering input ($\beta_1 RF + \beta_2 ER + \beta_3 VD$) only reaches 3 or 4 discrete values and the model works equally effectively with the perceptual inputs coded as discrete integers (e.g. ±1, ±2, ±3, with a different μ setting).

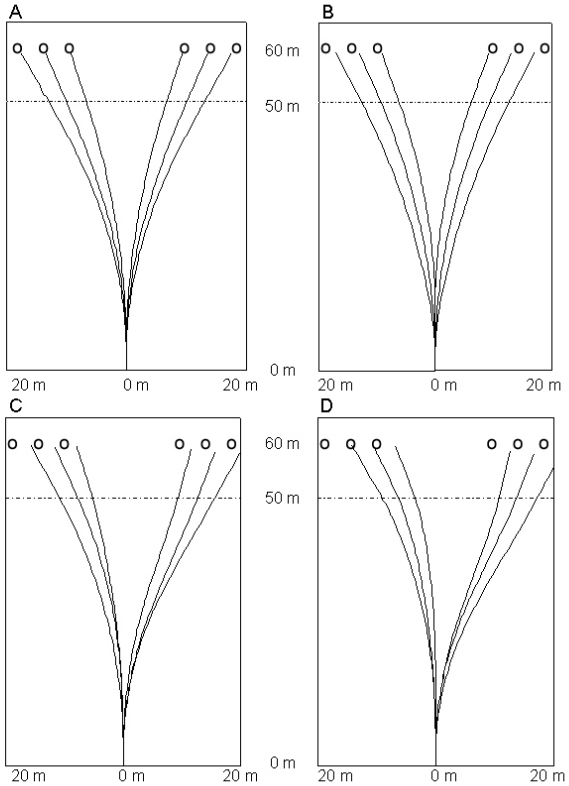

Figure A1. Response of a point attractor steering model to the conditions introduced in Experiment 3. Targets were placed 60 m ahead at 10, 14, and 18° to the left and right (small circles), speed of travel was 8 m/s, $\mu = 4.4$, $\upsilon = 1.8$ and $\beta_1 = \beta_2 = \beta_3 = 1/3$. A: Steering response with an observer sensitivity of 1°/s. B: Steering response with an observer sensitivity of 4°/s. C: Steering response with a retinal flow (RF) bias introduced alongside two other veridical sources (extraretinal, visual direction). D: Steering response with RF bias introduced alongside one other veridical source ($\beta_1 = \beta_2 = 1/2$, $\beta_3 = 0$).

Figure A1C illustrates the response of the model to the RF bias introduced in Experiment 3, when two other veridical sources of information (e.g., ER, VD) are available. Comparable to Experiment 3 the model was implemented with an additional ground rotation of bias of V*sin(20)/F until the 50 m point and then the path was extrapolated over the final 10 m to predict the steering error. The resultant errors from the model were 1.8 m for oversteer and –2.2 m for understeer responses.

Figure A1D repeats the simulation when only one additional source of information is available, comparable to the RR1V condition of Experiment 3, where VD information was absent. The predicted steering error in the latter case is 2.97 m for oversteer and –3.63 m for understeer (compare to Figure 11C).

Received February 14, 2001
Revision received May 2, 2002
Accepted August 27, 200

Journal of Experimental Psychology:
Human Perception and Performance
2003, Vol. 29, No. 2, 363-378