Pergamon

0042-6989(95)00122-0

Vision Res. Vol. 36, No. 3, pp. 431 443, 1996

0042-6989/96 $15.00 + 0.00

Estimating Heading During Real and Simulated Eye Movements

MARTIN S. BANKS,*† SHERYL M. EHRLICH,* BENJAMIN T. BACKUS,* JAMES A. CROWELL*

Received 20 October 1994; in revised form 19 April 1995

The ability to judge heading during tracking eye movements has recently been examined by several investigators. To assess the use of retinal-image and extra-retinal information in this task, the previous work has compared heading judgments with executed as opposed to simulated eye movements. For eye movement velocities greater than 1 deg/sec, observers seem to require the eye-velocity information provided by extra-retinal signals that accompany tracking eye movements. When those signals are not provided, such as with simulated eye movements, observers perceive their self-motion as curvilinear translation rather than the linear translation plus eye rotation being presented. The interpretation of the previous results is complicated, however, by the fact that the simulated eye movement condition may have created a conflict between two possible estimates of the heading: one based on extra-retinal solutions and the other based on retina-image solutions. In four experiments, we minimized this potential conflict by having observers judge heading in the presence of rotations consisting of mixtures of executed and simulated eye movements. The results showed that the heading is estimated more accurately when rotational flow is created by executed eye movements alone. In addition, the magnitude of errors in heading estimates is essentially proportional to the amount of rotational flow created by a simulated eye rotation (independent of the total magnitude of the rotational flow). The fact that error magnitude is proportional to the amount of simulated rotation suggests that the visual system attributes rotational flow unaccompanied by an eye movement to a displacement of the direction of translation in the direction of the simulated eye rotation.

Optic flow   Motion   Pursuit eye movement   Heading

INTRODUCTION

When a person walks through a scene while holding the direction of gaze fixed with respect to the direction of self-motion, the pattern of flow in the retinal image expands radially from a point—the focus of expansion—which is the projection of the person's heading (Gibson, 1966). However, when the person makes a tracking eye movement to fixate a moving object or a stationary object off the path, the pattern of retinal motion is no longer radial and the focus of expansion is obliterated (Regan & Beverley, 1982). Nevertheless, observers are still able to estimate their heading reasonably accurately (Royden, Banks & Crowell, 1992; Royden, Crowell & Banks, 1994; van den Berg, 1992; Warren & Hannon, 1988, 1990). To estimate heading accurately, observers must be able to estimate the translational component of the egomotion despite the confounding influence of rotational flow introduced by the rotation of the eye. How is this accomplished?

There are now numerous models of heading estimation in the presence of rotations. They fall into two categories: *retinal-image models* (e.g. Longuet-Higgins & Prazdny, 1980; Rieger & Lawton, 1985; Droulez & Cornilleau-Peres, 1990; Heeger & Jepson, 1992; Hildreth, 1992; Perrone & Stone, 1994) which employ retinal-image information only, and *extra-retinal models* (Royden *et al.*, 1994; von Holst, 1954; Wertheim, 1990) which make use of eye-velocity information from extra-retinal sources.

Although they use a variety of algorithms to estimate heading in the presence of rotational flow, all retinal-image models rely on the fact that flows due to translation and rotation have different properties. In particular, flow due to translation is depth-dependent, while flow due to rotation is not. One version of the retinal-image model is represented by the work of Longuet-Higgins and Prazdny (1980), Rieger and Lawton (1985), and Hildreth (1992); their algorithms subtract neighboring flow vectors and estimate heading by triangulation of the resulting difference vectors. This procedure works reasonably well because the rotational flow components are similar within neighborhoods, so the subtraction by and large eliminates their influence.

In contrast, extra-retinal models estimate the rotational components of retinal motion directly by means of extra-retinal signals; in the case of eye movements

*School of Optometry and Department of Psychology, University of California, Berkeley, CA 94720, U.S.A.
†To whom all correspondence should be addressed.

relative to the head, those signals could be provided by proprioceptive feedback from the extra-ocular muscles or by efferent signals to those muscles (von Holst, 1954; Matin, 1982). Assuming that rotational flow components can be determined accurately from those signals, it is a relatively simple matter to subtract the indicated rotational flow, compute the translational flow components, and then estimate the heading (Royden *et al.*, 1994).

The retinal-image approach is limited primarily by the visual input and the extra-retinal approach by the accuracy of non-retinal eye-velocity (or head-velocity) signals (Royden *et al.*, 1994). Therefore, a robust system would estimate heading both ways. When the two estimates differ, the system would choose (or weight more heavily) the one that generally provides more accurate estimates for the current viewing conditions.

A number of recent experiments have examined whether the visual system relies on one or both of these means of solution. In particular, Warren and Hannon (1988, 1990), Royden *et al.* (1992, 1994), and van den Berg (1992) presented the observer with two sorts of stimuli. In one condition—the *real eye movement* condition—observers judged their heading while making an eye movement to track a point in the simulated scene. In the second *simulated eye movement* condition, observers fixated a stationary point and the flow field deformed to simulate the effects of a tracking eye movement. To the degree that observers tracked the moving point accurately, the two conditions produced identical flow fields on the retina. Consequently, retinal-image models predict no difference in performance between the two conditions. Because the eyes did not move in the simulated condition, extra-retinal models predict misperceptions of heading in that condition.

Warren and Hannon (1988, 1990) reported that observers could discriminate headings equally well in the real and simulated eye movement conditions. From this, they concluded that human observers do not require extra-retinal information to judge heading in the presence of eye movements, an interpretation that supports the biological plausibility of retinal-image computational models. However, the eye movement velocities in those experiments were very slow at 0.2–0.7 deg/sec. These may be typical velocities for someone walking and tracking a distant object, but there are many situations in which people make much faster eye and head movements while moving through the world (Royden *et al.*, 1994). To explore the ability to solve the rotation problem in greater detail, Royden *et al.* (1992, 1994) measured human observers' heading judgments across a range of eye movement velocities. The first experiment reproduced Warren and Hannon's (1988) conditions except Royden and colleagues used constant, faster rotations of 0–5 deg/sec about a vertical axis. All observers responded quite differently in the real and simulated eye movement conditions; they judged heading accurately in the real movement condition regardless of rotation rate and very inaccurately in the simulated condition at rates greater than 1 deg/sec. In several subsequent experiments, Royden and colleagues showed that this outcome was not a consequence of the type of scene geometry or of the relationship between the fixation point and the scene. They concluded that human observers do in fact use extra-retinal information about eye position to judge heading accurately in the presence of rotations greater than about 1 deg/sec.

van den Berg (1992) showed that observers could estimate heading from noisy displays more accurately when rotational flow was caused by an executed rather than simulated eye movement. For noise-free displays, however, he reported an ability to perceive heading fairly accurately in the presence of simulated eye rotations of 0–5 deg/sec (van den Berg, 1993).

The results of Royden *et al.* (1992, 1994) do not allow us to conclude that observers rely entirely on extra-retinal, eye-velocity information to determine the rotational motion component; contributions of retinal-image information to estimating the rotation might have been masked in these experiments for two reasons. First, the stimuli in the experiments of Royden and co-workers were relatively small (40 × 40 deg or smaller) and retinal-image solutions should in principle be more accurate with larger fields of view (e.g. Koenderink & van Doorn, 1987). Second, the simulated eye movement condition differs from everyday situations in that it produces a potential cue conflict. In the real eye movement condition, the presumed eye-velocity signal matches the rotational flow in the retinal image (to the degree that it normally matches rotational flow; Mack & Herman, 1978). Thus, both methods of solving the rotation problem should yield similar estimates of heading in the real eye movement condition. With simulated eye movements, however, the eye-velocity signal provides information that the eye has not moved, yet the retinal image contains rotational flow that normally accompanies an eye movement. When this conflict becomes strong enough, as happens with fast simulated rotations, the visual system may reconcile the difference between the estimates of rotational motion provided by the two sources of information by suppressing the output of the retinal-image module (Royden *et al.*, 1994). For this reason, inaccurate judgments in the simulated condition cannot be taken as evidence that methods like the retinal-image models described above are not used in estimating heading under everyday conditions. Of course, it remains possible that retinal-image solutions are in fact not used at high rotation rates. The point is that the data of Royden and colleagues are not decisive on this issue.

In the work presented here, we examined the possibility that both solutions are employed in human vision and studied the manner in which they interact. There are several possible means of interaction. Here we describe three that differ according to how the two methods of solution are combined in determining the final heading estimate.

(1) *Winner takes all.* A heading estimate is derived from both methods and the observer's response reflects the estimate from the method that is deemed more

reliable for the current viewing situation. To explain the data of Royden and co-workers, one would have to assume that the extra-retinal solution determined the observers' choice in the presence of simulated eye rotations greater than 1 deg/sec.

(2) *Mixed.* The extra-retinal, eye-velocity information provides an estimate of the expected rotational flow. This estimate is subtracted from the observed flow and because the extra-retinal signal is not always completely accurate, the remainder of the rotational flow is estimated from retinal-image solutions. This model is consistent with the existing data.

(3) *Trigger.* The extra-retinal information provided by the execution of an eye movement provides one bit of information: whether the eye has moved or not. When it specifies that no eye rotation has occurred, the visual system seeks a solution to the observed flow that is consistent with no eye rotation (Royden *et al.*, 1994). When the signal specifies that an eye rotation has occurred, a solution is sought that includes a free parameter for the amount of eye rotation that may or may not correspond quantitatively with the actual rotation. This model is reasonably consistent with the existing data.

In the work presented here, we attempted to determine whether both means of solution are actually employed by human observers and, if so, how estimates from the two methods are weighted in the observers' reports of the direction of self-motion. We did so by presenting rotational flow composed of different proportions of simulated and executed eye rotation.

Predicted results are shown in Figs 1 and 2 for different ways in which extra-retinal and retinal-image solutions might interact. In Fig. 1, the error in perceived heading is plotted vs the rotation rate in the retinal image for a variety of simulated/real rotation mixtures. The upper left panel displays the predictions for the retinal-image model. Errors in perceived heading are small, are random rather than systematic, and are independent of the proportion of simulated/total rotation because the only determinant of performance is the rotation in the retinal image whether composed of real, simulated, or a mixture of real and simulated eye movement. The data of Royden *et al.* (1992, 1994) are inconsistent with these predictions. The upper right panel displays the predictions for the extra-retinal model. This model has the property that perceived heading errors should have a systematic component (a bias) that is proportional to the amount of simulated eye rotation—the amount of rotation in the retinal image that is not accompanied by an extra-retinal, eye-velocity signal. Thus, errors in perceived heading should be proportional to the magnitude of the total rotation and to the proportion of simulated/total eye movement; stated another way, the slopes of functions fit to the data should increase with increasing proportion of simulated/total eye movement. The lower left panel displays the predictions for the *mixed model* described above. Errors are small and unsystematic whenever the total rotation is small and when the rotation consists mostly of a real eye rotation. The lower right panel displays the predictions for the *trigger model* described above. Systematic errors occur with purely simulated eye rotations, but heading estimates are accurate whenever the eye actually rotated.

Figure 2 shows the same predictions when the error in perceived heading is plotted as a function of the amount of simulated eye rotation.

GENERAL METHODS

Three observers, all of whom were authors, participated in these experiments. They were corrected for the viewing distance with ophthalmic lenses.

The experiments were conducted on an Apple Macintosh IIfx computer with a 16-inch color monitor. The stimuli consisted of randomly-placed dots whose motions simulated various combinations of translation and rotation with respect to a ground plane or a 3D cloud; they yielded a sensation of self-motion. Each dot was one pixel subtending 6.9 by 6.9 min at the 18 cm viewing distance. The total angular subtense of the display varied from one experiment to the next (see

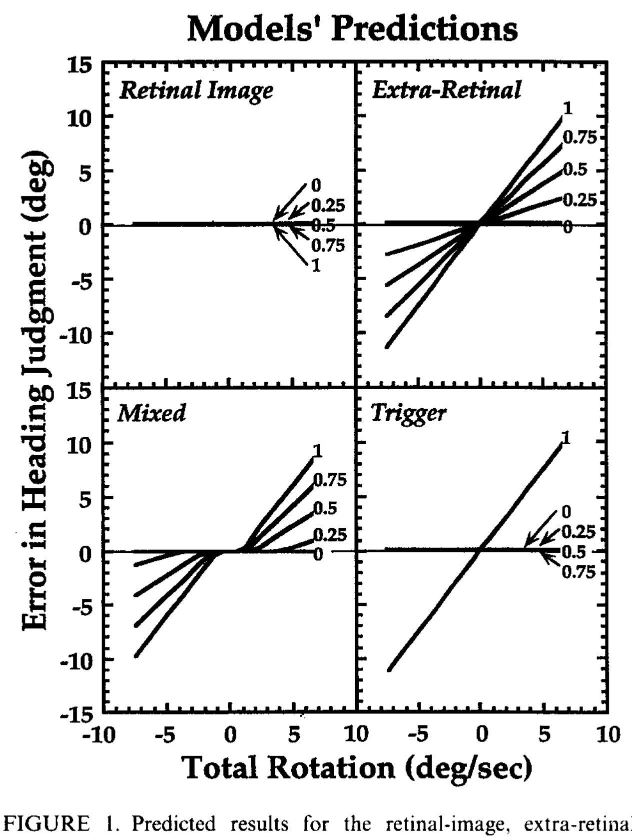

FIGURE 1. Predicted results for the retinal-image, extra-retinal, mixed, and trigger models (described in text). In each panel, the ordinate represents the predicted differences between observer responses and the depicted directions of self-motion (the plotted values are arbitrary). The abscissa in each panel represents the total rotational rate in the retinal image. The different sets of data show the predicted responses for different mixtures of simulated and real eye rotation. In the mixtures portrayed, the proportions of simulated eye rotation are 0.0, 0.25, 0.5, 0.75, and 1.0. A proportion of 0.0 corresponding to the condition in which the rotational flow is the consequence of a tracking eye movement. A proportion of 1.0 corresponds to one in which the observer's eye does not move and the rotational flow is added to the displayed motion sequence.

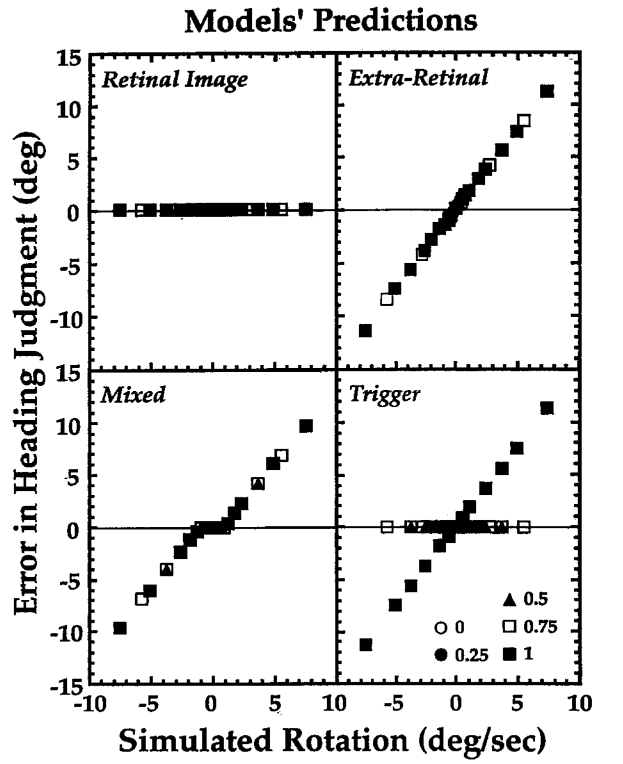

FIGURE 2. Model predictions plotted as a function of the amount of simulated eye rotation. The different symbols show the predicted responses for different proportions of simulated eye rotation.

later). Dot positions were updated at the 75 Hz frame rate. Head position was stabilized with a chin and forehead rest.

For these experiments, it is important to make environmental features including the edge of the display invisible. To achieve this, we viewed the stimuli monocularly through a 2 log unit neutral-density filter and an aperture. The room was completely dark except for the display. To ensure that observers would not be able to see the edge of the display screen during the course of the experiment, a bright uniform field was presented between trials to maintain light adaptation. With this setup, no environmental features could be seen throughout the experiment.

A fixation cross was provided and the observers were instructed to fixate and track it for the duration of each trial. The stimuli contained different proportions of simulated and real eye rotations. Specifically, for a given rotation rate, the proportion of the rotation due to a simulated eye movement was 0.0, 0.25, 0.5, 0.75, or 1.0. Proportions of 0.0 and 1.0 correspond respectively to the real and simulated eye movement conditions of Royden *et al.* (1992, 1994). For a proportion of 0.5, half of the total rotation was due to a smooth eye movement and half to a rotational flow component in the display.

Before each trial, the first frame of the forthcoming motion sequence appeared until the observer initiated the sequence with a button press. The fixation cross started moving 150 msec before the motion sequence so observers could establish smooth pursuit before the dots moved. At the end of a sequence, a cursor appeared and the observer positioned it to indicate perceived heading at trial end. No feedback was given. In a given experimental run, the mixtures, simulated headings, and rotation rates were randomly intermixed.

EXPERIMENT 1

Method

The stimulus in the first experiment simulated observer translation through a random 3D cloud of dots at initial distances of 18–500 cm. The simulated speed of translation was 150 cm/sec. The visible portion of the cloud subtended 30 × 30 deg and contained 64 dots at the beginning of the motion sequence; approx. 10 dots were visible at the end of the sequence. Stimulus duration was 1410 msec including 150 msec during which only the fixation cross moved. The headings portrayed by the motion sequences were all in the horizontal plane through the horizontal meridian of the display. Three headings were simulated: straight ahead, 3 deg to the left, and 3 deg to the right. The axis of all rotations (real and simulated) was vertical and the magnitudes were 0, 0.6, 1.25, 2.5, 5, and 7.5 deg/sec to the left and right. The proportions of simulated eye rotation were 0.0, 0.25, 0.5, 0.75, and 1.0 for all rotation magnitudes and headings. Although observers could place the cursor (indicating perceived heading) anywhere in the display, only the horizontal position was recorded.

Results and discussion

The results are displayed in Fig. 3; there is a separate panel for each of the three observers. Error in judged heading (averaged across the three headings) is plotted as a function of total rotation rate. The different lines in each panel represent the data for the various proportions of simulated eye rotation. Recall that a proportion of 0.0 corresponds to a condition in which the rotational flow in the retinal image is created entirely by an executed eye movement and a proportion of 1.0 corresponds to a condition in which the eye did not move and the rotational flow was entirely simulated. As reported by Royden *et al.* (1992, 1994), the magnitudes of the errors in heading judgments were proportional to the total rotation rate. More importantly, judgment errors also increased monotonically with increasing proportion of simulated eye rotation. For example, for observer MSB, the averages of the absolute values of the errors at ±7.5 deg/sec were 1.8, 3.0, 6.1, 7.6, and 10.8 deg for proportions of 0.0, 0.25, 0.5, 0.75, and 1.0, respectively.

Figure 4 illustrates that errors in judged heading depend strongly on the amount of simulated eye rotation. In this figure, heading error is plotted as a function of the magnitude of simulated eye rotation. Notice that all of the data could be fit well by one positively-sloped line regardless of the proportion of total eye rotation that was simulated. For example, consider the data at simulated rotation rates of 1.9–2.5 deg/sec. There are four data points at that rate corresponding to different simulated proportions. The total rotation rates for those points were 7.5 (0.25

simulated), 3.75 (0.5 simulated), and 2.5 deg/sec (0.75 and 1.0 simulated). Despite the three-fold variation in the magnitude of total rotational flow, the perceived heading errors were quite similar. Thus, the error in judged heading was coupled with the amount of simulated rotation and not with the amount of total rotation. This observation is consistent with the predictions of the extra-retinal model and not the retinal-image model. In addition, the observation offers little support for the hybrid models described in the Introduction.

The errors in perceived heading were large at high rotation rates and high simulated proportions. Observers in such cases perceived motion along a curved path while looking in the direction of instantaneous translation even though the motion sequence portrayed motion along a linear path while the direction of gaze changed. Given that observers misperceived the displays when the amount of simulated eye rotation was high, we wondered whether those errors were accompanied by a change in response variability. To examine this, Fig. 5 displays the standard deviations of responses for the conditions of this experiment. The standard deviations were lowest at total rotation rates near 0 deg/sec and increased monotonically at higher magnitudes of total rotation. Standard deviations did not vary systematically with the proportion of simulated rotation. Thus, observers were as consistent in making erroneous judgments at high proportions of simulated/total eye rotation as they were in making veridical judgments at low proportions.

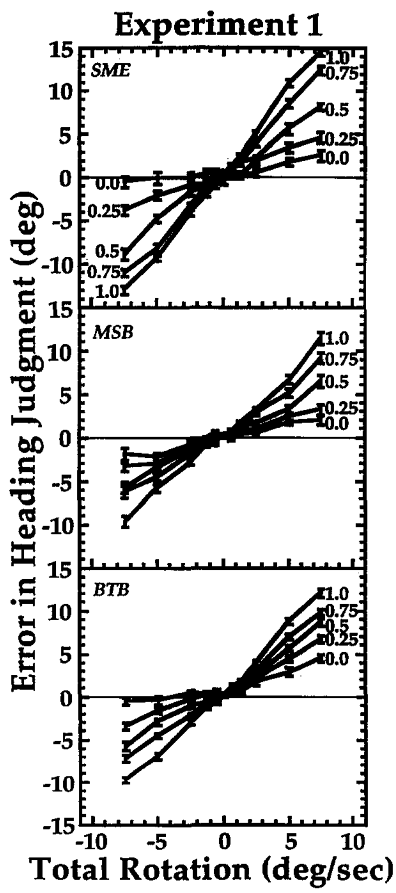

FIGURE 3. Heading judgment errors as a function of rotation rate for Experiment 1. The stimulus represented observer translation through a 3D cloud of dots. The ordinate represents the differences between the depicted directions of self-motion and the observers' responses; the displayed values are the averages of the differences across the three possible headings. The abscissa is the total rotational rate in the retinal image. The five sets of data in each panel represent judgments for different proportions of simulated eye rotation. If observers' judgments were accurate, the data would lie on the thin horizontal lines. Each data point represents the mean of 18–30 judgments (6–10 at each of three headings). The error bars represent ±1 SE.

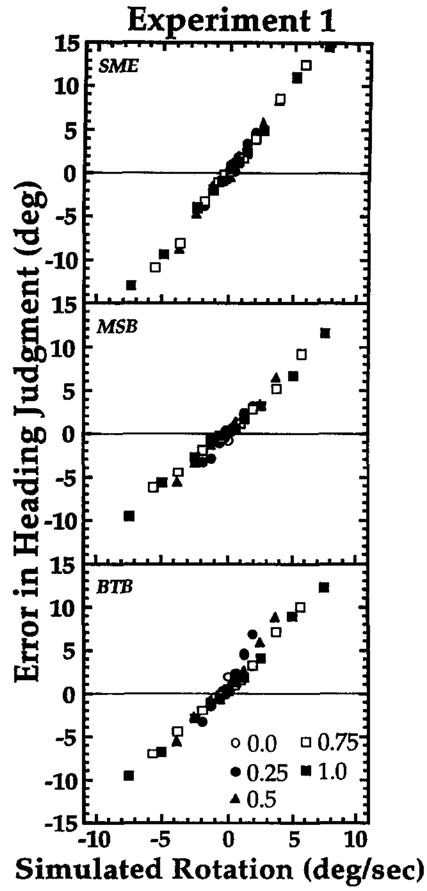

FIGURE 4. Heading judgment errors as a function of simulated rotation rate for Experiment 1. These are the same data that are plotted in Fig. 3, but the abscissa now represents the amount of simulated eye rotation rather than the amount of the total rotation. The five sets of data again represent judgments for different proportions of simulated eye rotation (indicated by the legend in the lower right). If observers' judgments were accurate, the data would lie on the thin horizontal lines.

EXPERIMENT 2

From a computational standpoint, the estimation of translation in the presence of simulated eye rotations should become more precise with an increasing field of view (e.g. Koenderink & van Doorn, 1987; Hildreth, 1992). Observers in Experiment 1 were unable to estimate heading accurately during large simulated eye rotations, so we next examined whether their inaccuracy was a consequence of the relatively small field of view.

Method

The stimulus and procedure of the second experiment were identical to the first with the exception that the stimulus subtended 60 × 40 deg and contained 128 dots at the beginning of the motion sequence (approx. 20 dots were visible at the end of the sequence). Two observers participated. SME was presented rotation rates of 0, 0.6, 1.25, 2.5, 5, and 7.5 deg/sec; in addition to those rates, MSB was presented 1.75 and 3.75 deg/sec.

Results and discussion

The results are displayed in Fig. 6. Again, the errors in perceived heading judgments increased with increases in the total rotation rate and in the proportion of simulated eye rotation. Figure 7 displays the same data plotted as a function of the magnitude of simulated eye rotation. Again, this method of plotting the data reveals that the errors in perceived heading were proportional to the magnitude of the simulated eye rotation. These large-field data are, therefore, consistent with the predictions of the extra-retinal model and inconsistent with the predictions of the others. As in Experiment 1, observers perceived a heading displaced from its true value in the direction of, and by an amount proportional to, the simulated eye rotation.

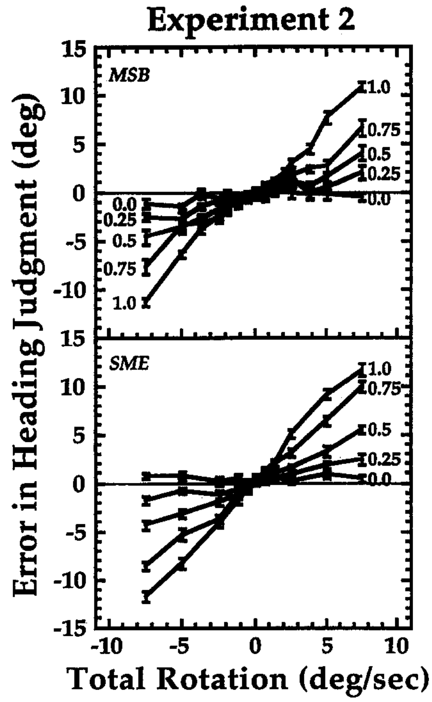

FIGURE 6. Heading judgment errors as a function of rotation rate for Experiment 2. As in Experiment 1, the stimulus depicted observer translation through a 3D cloud of dots, but the field of view was increased to 60 × 40 deg. The ordinate and abscissa are the same as in Fig. 3. Again, the five sets of data represent judgments for different proportions of simulated eye rotation. If observers' judgments were accurate the data would lie on the thin horizontal lines. Each data point represents the mean of 18–30 judgments (6–10 at each of three headings). The error bars represent ± 1 SE.

EXPERIMENT 3

There is some evidence that observers perceive heading more accurately during simulated eye rotations when the scene consists of a ground plane and the fixation point is attached to that plane (e.g. van den Berg, 1993; van den Berg & Brenner, 1994). The third experiment employed such a viewing situation.

Method

The stimulus depicted translation parallel to and 160 cm above a ground plane. The plane had a simulated depth of 4000 cm and 150 dots were visible on average. Field of view was 64 × 64 deg. The speed of translation was 190 cm/sec and the directions were straight ahead, 4 deg to the left, and 4 deg to the right. Trial duration was 1006 msec. The fixation point was a point on the ground plane, so the rotations were about various axes. The horizontal components of the rotations (the average values for the duration of a motion sequence) were 0, 0.5, 1.0, 1.8, 2.5, and 5 deg/sec to the left; observer MSB was not presented the rotations of 0.5 and 1.8 deg/sec. At the end of each motion sequence, the last frame was displayed along with a horizontal line coincident with the projection of the horizon. Observers indicated the perceived heading by placing the cursor along this line. Five proportions of simulated eye rotation were presented: 0.0, 0.25, 0.5, 0.75, and 1.0.

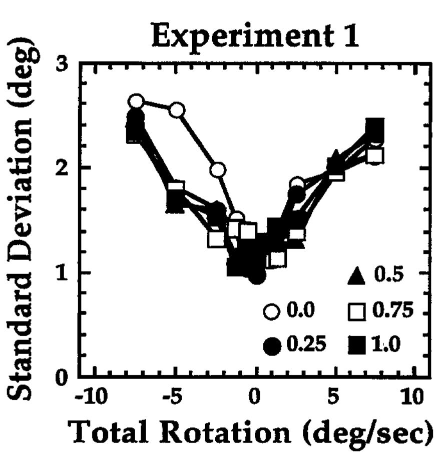

FIGURE 5. Variability of heading judgments as a function of rotation rate for Experiment 1. The ordinate represents the standard deviations of observer responses; the displayed values are the averages of the differences across the three observers and the three possible headings. The abscissa is the total rotational rate in the retinal image. The five sets of data represent judgments for different proportions of simulated eye rotation (indicated by the legend in the lower right).

Results and discussion

The results are displayed in Fig. 8. Because the rotations in this experiment were all in the same direction, the format of this figure differs from the figures showing the model predictions (Figs 1 and 2) and the 3D cloud data (Figs 3, 4, 6, and 7). Specifically, Figs 8 and 9 show data in the upper right quadrant of the previous figures.

The three observers behaved differently, but the errors in their heading judgments increased with increasing total rotation and with increasing proportion of simulated eye rotation. Observer MSB did not exhibit large errors for any of the conditions. Observers SME and BTB exhibited large errors under some conditions. For SME, errors were essentially proportional to the proportion of simulated eye rotation. For BTB, errors were not proportional to the proportion of simulated rotation because his responses were rather inaccurate when the simulated proportion was small. This observer had difficulty with the task because the ground plane around the fixation point appeared non-rigid to him. He performed the experiment again after being instructed to attend to the entire field of view; his responses were generally more accurate at low proportions of simulated rotation and less accurate at high proportions, but the pattern of results still did not closely resemble those of observer SME or MSB.

Figure 9 displays the same data plotted as a function of the magnitude of simulated eye rotation. Again, it is clear that the three observers behaved differently in this experiment. The heading judgments of observer SME were proportional to the proportion of simulated eye rotation. The judgments of observer MSB were reasonably accurate for all conditions. Those of observer BTB increased with increasing proportion of simulated rotation, but his responses were inaccurate at low rotation rates.

Because the three observers behaved differently, one cannot draw strong conclusions from the results of this experiment. Observer MSB exhibited a pattern of results that was reasonably consistent with the retinal-image model; as predicted by that model, his heading judgments were reasonably accurate at all rotation rates and all mixtures of simulated and real eye rotations. In contrast, the data from observer SME were clearly inconsistent with the retinal-image predictions

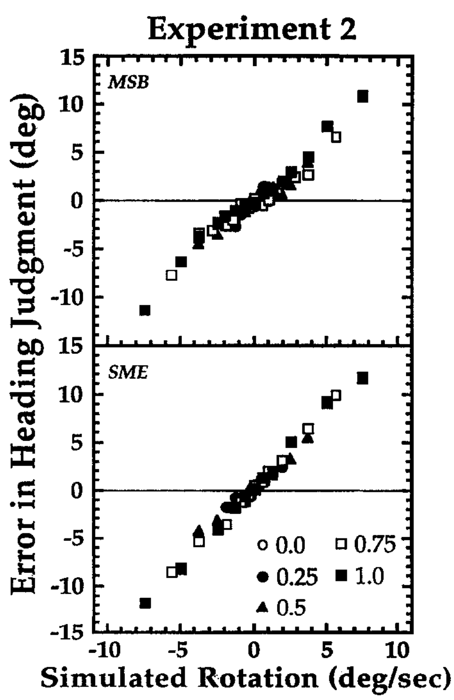

FIGURE 7. Heading judgment errors as a function of simulated rotation rate for Experiment 2. These are the same data plotted in Fig. 6, but the abscissa now represents the amount of simulated eye rotation rather than the amount of the total rotation. The five sets of data represent judgments for different proportions of simulated eye rotation. If observers' judgments were accurate, the data would lie on the thin horizontal lines.

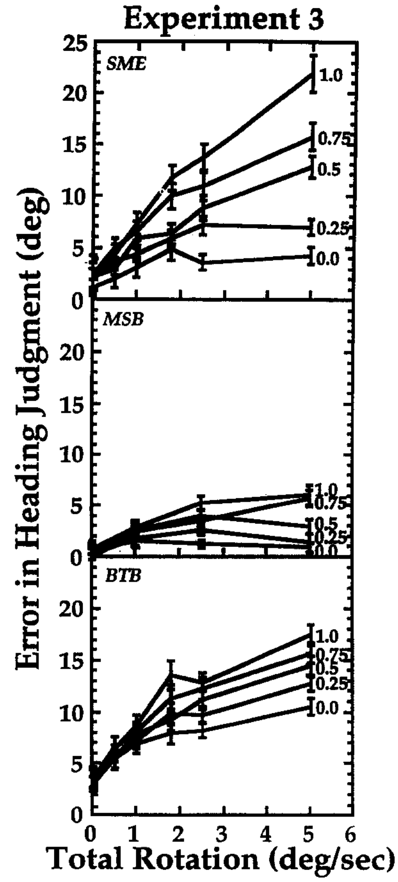

FIGURE 8. Heading judgment errors as a function of rotation rate for Experiment 3. The stimulus depicted observer translation parallel to a ground plane; observers responded by positioning a cursor along a horizontal line coincident with the horizon. The ordinate again represents the differences between the depicted directions of self-motion and the observers' responses (averaged across the three possible headings). The abscissa is the total rotational rate in the retinal image; the signs were all the same because the rotations were always leftward and downward. The five data sets represent judgments for different proportions of simulated rotation. If observers' judgments were accurate, the data would lie on the abscissa. Each data point represents the mean of 18–30 judgments (6–10 at each of three headings). Error bars represent ± 1 SE.

and were most consistent with the those of the extra-retinal model. Observer BTB had difficulty with the task and his data are not consistent with any of the models' predictions.

EXPERIMENT 4

Translation parallel to a ground plane with a simulated eye rotation provides a simple cue that is not present in most other situations (van den Berg, 1992; Royden *et al.*, 1994). In this situation, heading corresponds with the intersection of the horizon and a line containing points with common flow directions. Royden *et al.* (1994) provided some evidence that observers can capitalize on this cue to estimate heading during simulated eye rotations. Observer MSB reported using this cue in Experiment 3, but SME and BTB were not aware of using it. We examined the possibility that observers, particularly MSB, used this so-called horizon cue by eliminating it in a fourth experiment.

Method

The stimulus depicted translation relative to a ground plane. In order to eliminate the horizon cue, the translation contained a variable component perpendicular to the ground plane. Specifically, the headings lay along a 45 deg diagonal line on the display screen; the line intersected the midpoint of the screen and increased in height from left to right. The fixation point was always positioned on this line, so rotations were always about an axis tilted 45 deg from vertical. Three headings were presented: straight ahead, 4 deg down and to the left of straight ahead, and 4 deg up and to the right. Observers indicated perceived headings by positioning a cursor along the diagonal line. In other respects, the displays were identical to those in Experiment 3.

There was a cue in the experiments of Warren and Hannon (1988, 1990) and some of the experiments of Royden *et al.* (1992, 1994) that may have confounded comparisons of performance during real and simulated eye rotations. In the real eye movement condition, the fixation point moved relative to the boundaries of the displayed flow field and in the simulated condition, it did not. Because motion of the fixation target with respect to the surroundings normally accompanies eye movements, observers could have used the motion of the fixation point relative to the edge of the visible display in the real movement condition to estimate the velocity of the rotational flow component. We eliminated this difference between real and simulated eye rotation displays by shifting the software clipping window in the direction and by the amount of the real eye rotation. The observers did not notice this difference in the displays.

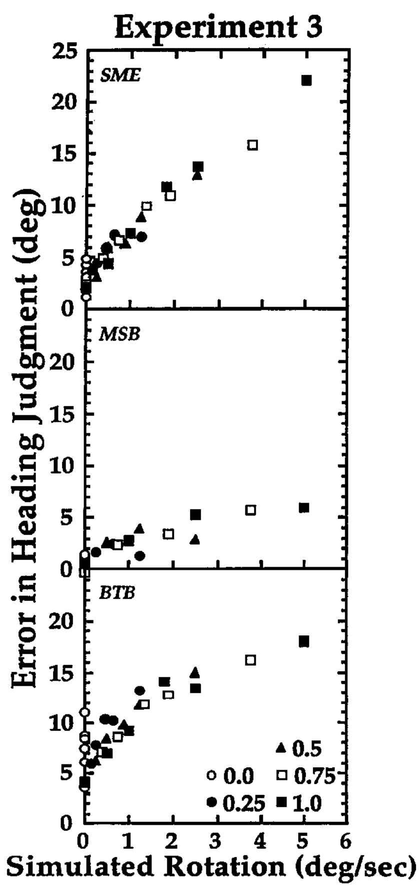

FIGURE 9. Heading judgment errors as a function of simulated rotation rate for Experiment 3. These are the same data plotted in Fig. 8, but the abscissa now represents the amount of simulated rotation. The five sets of data represent judgments for different proportions of simulated eye rotation (as indicated by the legend in the lower right). If observers' judgments were accurate, the data would lie on the abscissa.

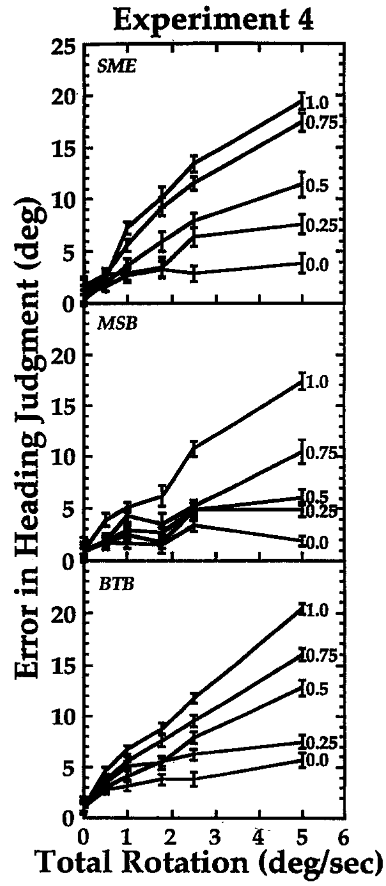

FIGURE 10. Heading judgment errors as a function of rotation rate for Experiment 4. As in Experiment 3, the stimulus depicted observer translation relative to a ground plane, but now the translation had a variable vertical component and the fixation point was always positioned near a diagonal line that intersected the horizon at the midpoint of the screen. Observers responded by positioning a cursor along the diagonal line. The ordinate again represents the differences between the depicted directions of self-motion and the observers' responses (averaged across the three possible headings). The abscissa is the total rotational rate in the retinal image; the signs were all the same because the rotations were always leftward and downward. The five data sets represent judgments for different proportions of simulated rotation. If observers' judgments were accurate, the data would lie on the abscissa. Each data point represents the mean of 18–30 judgments (6–10 at each of three headings). Error bars represent ± 1 SE.

Results and discussion

The results are displayed in Fig. 10. Unlike the outcome of Experiment 3, the three observers behaved quite similarly. Errors in heading judgments were proportional to the total rotation and to the amount of simulated rotation. All three observers reported perceiving a curvilinear path of motion toward the fixation point when the proportion of simulated/total rotation was high. Observer BTB, who had difficulty with the task in Experiment 3, reported less difficulty in the fourth experiment. Observer MSB, whose judgments were fairly accurate under all conditions in Experiment 3, perceived heading less accurately at high proportions of simulated eye rotation. Therefore, elimination of the horizon cue had no effect on observer SME, led to poorer performance by MSB, and produced a different pattern of results in observer BTB.

Figure 11 displays the same data plotted as a function of the magnitude of simulated eye rotation. This plotting method reveals that the errors in perceived heading are proportional to the magnitude of simulated eye rotation. These ground-plane data are, therefore, consistent with the predictions of the extra-retinal model.

DISCUSSION

Implications for models of heading estimation

The results of Experiments 1, 2, and 4 were similar across experiments and observers. The results of Experiment 3 differed across observers, so they will be discussed in a later section.

As observed by Rieger and Toet (1985), Royden and colleagues (1992, 1994), and van den Berg (1992), we found in Experiments 1, 2, and 4 that observers make inaccurate heading judgments in the presence of rotational flow that is unaccompanied by an eye movement. Additionally, we found that the magnitudes of errors in such judgments are essentially proportional to the amount of simulated rotation; this observation is best illustrated by Figs 4, 7, and 11 which plot error magnitude as a function of simulated rotation rate. These figures show that error magnitude was essentially a linear function of simulated rotation rate regardless of the total rotation rate in the retinal image. All three observers perceived headings displaced in the direction of, and by an amount proportional to, the simulated eye rotation. In particular, they reported perceiving a curved motion path in the direction of the simulated rotation (for an interpretation of this phenomenon, see Royden, 1994).

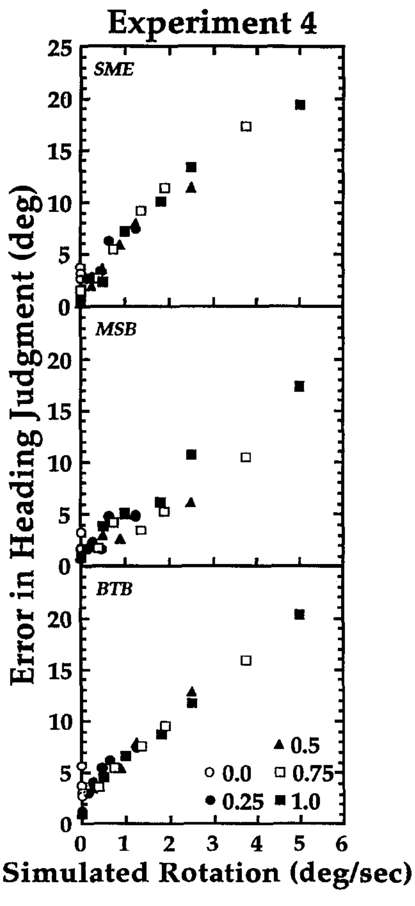

FIGURE 11. Heading judgment errors as a function of simulated rotation rate for Experiment 4. These as the same data plotted in Fig. 10, but the abscissa now represents the amount of simulated rotation. The five sets of data represent judgments for different proportions of simulated eye rotation (as indicated by the legend in the lower right). If observers' judgments were accurate, the data would lie on the abscissa.

It is obvious from a comparison of the data figures (3, 4, 6, 7, 10, and 11) and the model predictions shown in Figs 1 and 2 that observers' heading judgments were inconsistent with the retinal-image and trigger models described earlier. Both of those models predict that judgment errors should not be affected by the proportion of simulated/total eye rotation (except at high proportions for the trigger model), but judgments were clearly affected by this proportion.

Observers' judgments were also generally inconsistent with the mixed model as presented in Figs 1 and 2. According to this model, heading errors occur whenever the difference between the rotational flow in the retinal image and an extra-retinal, eye-velocity signal exceeds a criterion value. In particular, heading judgments should not vary with the proportion of simulated eye rotation at slow total rotation rates. The data of Figs 3, 4, 6, 7, 10, and 11 generally disconfirm this prediction; within the measurement error limits of these experiments, there appears to be no such region of independence. We note, however, that some of the data from observer MSB (Fig. 6) are reasonably consistent with the mixed model; in particular, at rotation rates of 0.6 and 1.25 deg/sec, his heading judgments did not vary significantly with mixture. MSB is by far the most experienced of the three observers, so we speculate that his ability to judge heading at slow simulated rotation rates is a consequence of greater familiarity with the stimuli and task.

Our observations are most consistent with the predictions of the extra-retinal model described in the Introduction. As predicted by that model, errors were essentially proportional to the ratio of simulated/total eye rotation regardless of the total rotation rate. At least for the conditions of Experiments 1, 2, and 4, it seems that rotational flow that is incommensurate with the velocity of an observer's tracking eye movement is attributed to a displacement of the direction of translation rather than to a rotation of the eye with respect to the simulated scene.

Royden (1994) has interpreted the displacement of perceived heading as follows. The extra-retinal, eye-velocity signal produced while tracking a point provides an estimate of the rotational flow. This estimate is then subtracted from the observed flow and the remaining rotational flow is attributed to a curvature in the observer's path of motion. This idea is plausible because the retinal flow field caused by curvilinear translation while holding the eye fixed (relative to the head) is similar to the flow field caused by linear translation while rotating the eye (Warren, Mestre, Blackwell & Morris, 1991).

Heading estimates in the ground plane experiments

The results of the two ground plane experiments (3 and 4) differed from one another. In Experiment 4, all observers exhibited a similar pattern of perceived heading errors, but this was not the case in Experiment 3. Specifically, one observer (SME) in Experiment 3 exhibited errors very much like those in the other experiments, another (MSB) was able to estimate heading in the presence of simulated rotation more accurately in this experiment than in the others, and another observer (BTB) had difficulty performing the task. We have no straightforward explanation for why the same observers behaved dissimilarly when translation was parallel to a ground plane (Experiment 3) but similarly when a vertical translation component was added (Experiment 4). It is interesting to note, however, that the display producing the most variable behavior—translation parallel to a ground plane while fixating a point in the plane—is perhaps the most commonly used display in this area of research (Royden *et al.*, 1994; van den Berg, 1992, 1993; van den Berg & Brenner, 1994; Warren & Hannon, 1988, 1990).

With the exception of observer MSB, the data of Experiment 3 were inconsistent with the retinal-image and trigger models described earlier. These models predict that errors in heading judgments should not vary with the mixture of real and simulated eye rotation, but errors did in fact vary with the mixture. The data of observer MSB, however, were reasonably consistent with the predictions of the retinal-image model because his judgments were reasonably accurate for all mixtures of simulated and real eye rotation.

Comparison with previous reports

The data from Experiments 1, 2, and 4 are generally consistent with the existing literature. For the scenes and observer motions portrayed in those experiments, observers perceived heading reasonably accurately at slow rotation rates whether the rotational flow was due to an executed or simulated eye movement; this observation is similar to those of Royden *et al.* (1992, 1994) and Warren and Hannon (1988, 1990). It is important to note, however, that most of the heading judgments at slow rotation rates (1.25 deg/sec or slower) were in fact affected by the mixture of simulated and real eye rotation; in particular, they were more accurate when the proportion of simulated rotation was small; the exception was observer MSB's judgments in Experiment 2 (Fig. 6). Our finding that heading was misperceived at high rotation rates (>1.25 deg/sec) when the rotational flow was due to a simulated eye movement is consistent with the observations of Rieger and Toet (1985) and Royden *et al.* (1992, 1994).

The latter observation seems inconsistent with recent reports by van den Berg (1993) and van den Berg and Brenner (1994) who reported accurate heading perception during simulated eye rotations as high as 5 deg/sec. We wish to make two points in regard to the van den Berg data.

First, the differences between their data and ours are not consistent. Indeed, some of the displayed results in van den Berg (1993) and van den Berg and Brenner (1994) are actually very similar to our findings. There are two figures in those reports with sufficient data to allow a comparison of their results and ours: Fig. 1 in van den Berg (1993) and Fig. 2 in van den Berg and Brenner (1994). In both cases, the authors plotted "perceived heading" as a function of "simulated heading". The former refers to the horizontal angle at trial end between the fixation point and the position of the observer's response; the latter refers to the horizontal angle at trial end between the fixation point and the position of the simulated heading. The van den Berg data have been replotted in Fig. 12 in the format used here along with data from Experiments 2 and 3. The left panel displays van den Berg and Brenner's (1994) data for a 3D cloud presented stereoscopically and data from observers MSB and SME in Experiment 2; they are clearly quite similar. In other words, this observer exhibited errors similar to the ones we observed when rotation was entirely simulated. Apparently, some of the other observers in van den Berg and Brenner (1994) gave more accurate responses than the one shown here. Specifically, they plotted the slopes of the response vs stimulus functions in their Fig. 3 and the other two observers yielded greater slopes than the one whose data are plotted in Fig. 12. Figure 12(b) displays ground plane data from van den Berg and Brenner's (1994) (stimulus presented stereoscopically), ground plane data from van den Berg's (1993) (stimulus presented monocularly), and ground plane data from observers MSB, SME, and BTB in Experiment 3. In this case, the results are noticeably more variable; van den Berg and Brenner's (1994) observer gave heading responses similar to MSB, but van den Berg's (1993) observer gave more accurate responses. In summary, some but not all of the data from the van den Berg reports are actually quite similar to the data reported here. Where the data are different, however, van den Berg's observers gave more accurate responses.

The second point concerning the comparison with van den Berg's data concerns the experience of the observers. The stimuli in van den Berg (1993) all depicted translation parallel to a ground plane while fixating a point in that plane; this stimulus was one of the two employed by van den Berg and Brenner (1994). This display is quite similar to the one presented in Experiment 3 and it is interesting to note that the most experienced of our three observers behaved in much the same fashion as van den Berg's observers: MSB was able to estimate heading fairly accurately during simulated eye rotations as high

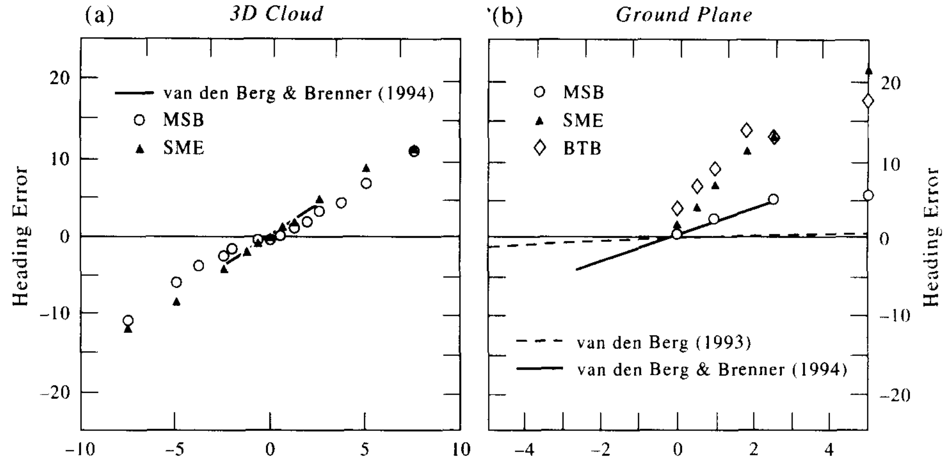

FIGURE 12. Comparison of data from van den Berg (1993) and van den Berg and Brenner (1994) with the data of Experiments 2 and 3. The error in perceived heading (in degrees) is plotted as a function of the simulated rotation rate. For the data of Experiments 2 and 3, the plotted values are drawn from trials in which the rotations were 100% simulated. (a) Results for a 3D cloud stimulus. In Experiment 2, the cloud was presented monocularly; in the van den Berg and Brenner (1994) paper, the cloud was presented stereoscopically. The thick solid line represents the van den Berg and Brenner data; it was calculated from the regression line plotted in the right panel of their Fig. 2. The symbols represent average errors in perceived heading from Experiment 2 of the current manuscript; the data from the two observers are represented by different symbols. (b) Results for a ground plane stimulus. The thick solid line represents the van den Berg and Brenner (1994) data; it was calculated from the regression line shown in the left panel of their Fig. 2. The dashed line represents the data of van den Berg (1993); it was calculated from the regression line shown in Fig. 1 of that paper.

as 5 deg/sec. The observers in van den Berg (1993) and van den Berg and Brenner (1994) were also quite experienced with optic flow displays, so very experienced observers may be able to estimate heading during simulated eye rotations fairly accurately. The specific conditions under which heading can be estimated in the presence of a simulated tracking eye movement remain to be delineated, but the observer's experience may be an important factor.

Perceiving curvilinear translation during simulated rotations

Heading judgments in these experiments were generally displaced in the direction of the rotation by an amount roughly proportional to the amount of simulated rotation (see also Royden *et al.*, 1992, 1994). Observers frequently perceived a curvilinear path of self-motion rather than the linear translation plus rotation that was being simulated; this misperception was most apparent in the ground plane experiments. Why would observers perceive curvilinear motion in these displays? As pointed out previously (Warren *et al.*, 1991), movement along on a circular path can be described *instantaneously* as the sum of a translation along the tangent to the path and an eye rotation about a perpendicular axis. Consequently, the instantaneous flow field for forward translation plus an eye rotation about a vertical axis is identical to the field for curvilinear motion about a vertical axis with eye position fixed (with respect to the body and head). The retinal image motions induced by these two situations differ more and more as time passes, so one's ability to distinguish the two situations might well depend on the duration of the stimulus and other stimulus properties such as the speed of translation and the field of view.

Given the similarity of the flow fields for curvilinear translation and linear translation plus rotation, it is important to ask how different the fields would be for the conditions of our experiments. Figure 13 depicts the retinal image motions associated with the two situations for the conditions of Experiments 2 and 3; (a) shows the motions for the ground plane stimulus of Experiment 3 and (b) shows the motions for the cloud stimulus of Experiment 2. In both panels, the black lines represent the retinal trajectories of a subset of dots for linear translation plus a rotation of 5 or 7.5 deg/sec. The gray lines represent the trajectories of the same dots for a curvilinear translation; the radius of curvature for this translation was chosen to render the flow field similar to the one for translation plus rotation. The retinal image motions are indeed fairly similar, but the separations between trajectories become larger with time and are larger for near dots than for distant dots. At trial end, the largest separations between the sets of trajectories are 1.5 deg for the ground plane stimulus (lower left part of the field) and 2.2 deg for the 3D cloud stimulus. Those separations and the associated changes in vector velocity are easily discriminable even in the peripheral visual field when the targets are single dots or lines (McKee & Nakayama, 1984). Thus, observers should be able to *discriminate* the flow fields associated with curvilinear translation from those associated with linear translation plus rotation. Nonetheless, observers seem to erroneously perceive a curvilinear path of motion in the presence of a simulated eye rotation presumably because extra-retinal signals imply that the eye has not rotated. We assume that such misperceptions do not occur in everyday settings with observer-initiated locomotion because extra-retinal information from the extra-ocular muscles, neck muscles, and the vestibular system helps

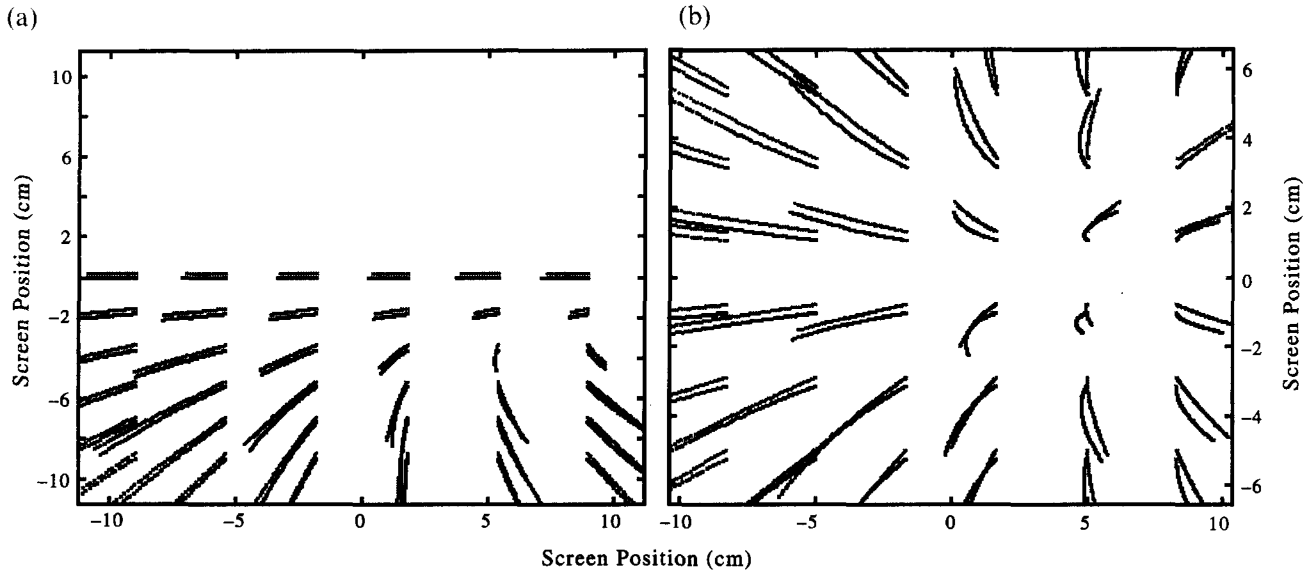

FIGURE 13. Typical dot trajectories across the display screen. (a) The trajectories for the dots in Experiment 3. (b) The trajectories in Experiment 2. The black points represent trajectories for linear translation plus rotation (i.e. the simulated eye rotation condition). The gray points represent trajectories for curvilinear translation with gaze always in the instantaneous direction of translation. The gray points have been shifted slightly vertically relative to the black points. The two sets of dot paths are similar. Parameters for (a): viewing distance = 18 cm, field of view = 64 × 64 deg, duration = 1006 msec, observer speed = 190 cm/sec, and simulated eye-height = 160 cm. For the linear translation plus rotation, rotation rate (R_y) = 5 deg/sec; for the curvilinear translation, R_y = 4.5 deg/sec; this value of R_y was chosen to provide the smallest overall separation between the flow vectors in the two situations. Parameters for (b): viewing distance = 18 cm, field of view = 60 × 40 deg, duration = 1260 msec, observer speed = 150 cm/sec, and initial dot depths = 18–500 cm. For the linear translation plus rotation, rotation rate (R_y) = 7.5 deg/sec; for the curvilinear translation, R_y = 5.25 deg/sec; again, this value of R_y was chosen to provide the smallest overall separation between the flow vectors in the two situations.

distinguish linear self-motion during an eye/head rotation from motion along a curved path with the eye and head fixed.

CONCLUSION

The potential conflict created by the simulated eye movement condition of previous experiments was minimized by having observers judge heading in the presence of rotations consisting of mixtures of executed and simulated eye movements. The results of Experiment 3 were difficult to interpret because the three observers behaved differently. The observers responded very similarly in Experiments 1, 2, and 4. In those experiments, heading was estimated more accurately when rotational flow was created by executed eye movements alone. Most importantly, the magnitudes of errors in heading estimates were essentially proportional to the amount of rotational flow created by a simulated eye rotation. The fact that error magnitudes were proportional to the amount of simulated rotation suggests that the visual system attributes rotational flow unaccompanied by an eye movement to a displacement of the observer's path in the direction of and by an amount proportional to the simulated eye rotation. There was some evidence that a highly experienced observer can estimate heading during simulated eye rotations under some conditions.

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. (1993). Perception of heading. *Nature*, *365*, 497–498.

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

Droulez, J. & Cornilleau-Peres, V. (1990). Visual perception of surface curvature. The spin variation and its physiological implications. *Biological Cybernetics*, *62*, 211–224.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton-Mifflin.

Heeger, D. J. & Jepson, A. D. (1992). Subspace methods for recovering rigid motion I: Algorithm and implementation. *International Journal of Computer Vision*, *7*, 95–117.

Hildreth, E. C. (1992). Recovering heading for visually-guided navigation. *Vision Research*, *32*, 1177–1192.

von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behavior*, *2*, 89–94.

Koenderink, J. J. & van Doorn, A. J. (1987). Facts on optic flow. *Biological Cybernetics*, *56*, 247–254.

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

Mack, A. & Herman, E. (1978). Loss of position constancy during pursuit eye movements. *Vision Research*, *18*, 55–62.

Matin, L. (1982). Visual localization and eye movements. In Wertheim, A. H., Wagenaar, W. A. & Leibowitz, H. W. (Eds), *Tutorials on motion perception*. New York: Plenum Press.

McKee, S. P. & Nakayama, K. (1984). The detection of motion in the peripheral visual field. *Vision Research*, *24*, 25–32.

Perrone, J. A. & Stone, L. S. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision Research*, *34*, 2917–2938.

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.

Rieger, J. H. & Lawton, D. T. (1985). Processing differential image motion. *Journal of the Optical Society of America A*, *2*, 354–360.

Rieger, J. H. & Toet, L. (1985). Human visual navigation in the presence of 3D rotations. *Biological Cybernetics*, *52*, 377–381.

Royden, C. S. (1994). Analysis of misperceived observer motion during simulated eye movements. *Vision Research*, *34*, 3215–3222.

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

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

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 A*, *7*, 160–169.

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.

Wertheim, A. H. (1990). In Warren, R. & Wertheim, A. H. (Eds), *Perception and control of self-motion*. Hillsdale, N.J.: Earlbaum.

Acknowledgements—We thank Kirk Swenson for software assistance and Bill Warren for helpful discussions. This research was supported by AFOSR Research Grant 93NL366, NSF Research Grant DBS-9309820, and NIH Training Grant EY-07043.