

Seeking one's heading through eye movements

JAMES E. CUTTING

Cornell University, Ithaca, New York

PAULA M. Z. ALLIPRANDINI

State University of São Paulo, São Paulo, Brazil

and

RANXIAO FRANCES WANG

University of Illinois, Urbana, Illinois

A study of eye movements during simulated travel toward a grove of four stationary trees revealed that observers looked most at pairs of trees that converged or decelerated apart. Such pairs specify that one's direction of travel, called *heading*, is to the outside of the near member of the pair. Observers looked at these trees more than those that accelerated apart; such pairs do not offer trustworthy heading information. Observers also looked at gaps between trees less often when they converged or diverged apart, and heading can never be between such pairs. Heading responses were in accord with eye movements. In general, if observers responded accurately, they had looked at trees that converged or decelerated apart; if they were inaccurate, they had not. Results support the notion that observers seek out their heading through eye movements, saccading to and fixating on the most informative locations in the field of view.

Our eyes rove as we walk. Only 10% of the time do we look within 5° of our direction of movement (Wagner, Baird, & Barbaresi, 1981), or *heading*. Yet, as pedestrians, we need to know where we are going within 3.5° to avoid accidents (Cutting, Springer, Braren, & Johnson, 1992). How is it we can walk about and find our way safely? Many computational schemes have been offered (e.g., Hildreth, 1992; Perrone & Stone, 1994; Rieger & Lawton, 1985). These focus on pooling motions over large regions of the visual field, and neurophysiological evidence supports their feasibility (Bradley, Maxwell, Andersen, Banks, & Shenoy, 1996; Duffy & Wurtz, 1991). Such schemes, however, do not always predict human heading judgments in naturalistic environments (Cutting, 1996; Cutting, Wang, Flückiger, & Baumberger, 1999). Instead, for guidance, we propose that pedestrians pay attention to objects, their relative depth, and their relative motion, ignoring other motion in the visual field.

A moving observer is immersed in a sea of motion that can be parsed in many ways, fostering many different approaches to wayfinding (see Warren, 1995, 1998, for reviews). Here, we pose a new approach. Consider all stationary objects in the forward visual field. There are three

possible pairwise relative motions, regardless of where one looks: Objects can converge, they can decelerate apart, or they can accelerate apart (Cutting, 1996; Wang & Cutting, 1999). Figure 1A shows a plan view of a pedestrian and a given object off his or her path to the right, as well as the areas occupied by second objects creating pairs that converge, decelerate apart, and accelerate apart. Since these motion classes stem from the objects' locations relative to each other and to the observer's instantaneous direction of translation, they apply equally to straight and curved paths (Wang & Cutting, 1999). Figure 1B shows the statistical probability of the location of one's heading in the three instantaneous cases, where arrows indicate the relative directions of motion, and the diagonals indicate the lines of sight to each tree, parsing the visual field into three response categories: to the outside of the near tree, between the trees, or to the outside of the far tree.

The first two situations are straightforward: Convergence and decelerating divergence between two stationary objects specify that one's heading direction is always outside the near object. In these examples, it is to the left. Heading never lies between the two trees, nor outside the far object (in this case, to the right). After Gibson (1979), we call these *invariants* because of their trustworthy nature, and they are invariant under the transformation of the observer's forward movement. Notice, however, that they are nominal invariants; each indicates only that one's heading direction is to the left or right of a given object, but not how far. Nonetheless, Wang and Cutting (1999) found that, with two nominal invariants of opposing sign (one on either side of the direction of movement), observers' judgments were constrained to within 0.8° of the heading.

J.E.C. was supported by NSF Grant ASC-9523483, and P.M.Z.A. was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo Grant 97/01119-7. The authors thank Evan Creutz for his help (particularly in the continuous calibration of the eye tracker), Nan Karwan and Michael Spivey for discussion, and Bruce Halpern for comments. Correspondence should be addressed to J. E. Cutting, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853-7601 (e-mail: jec7@cornell.edu).

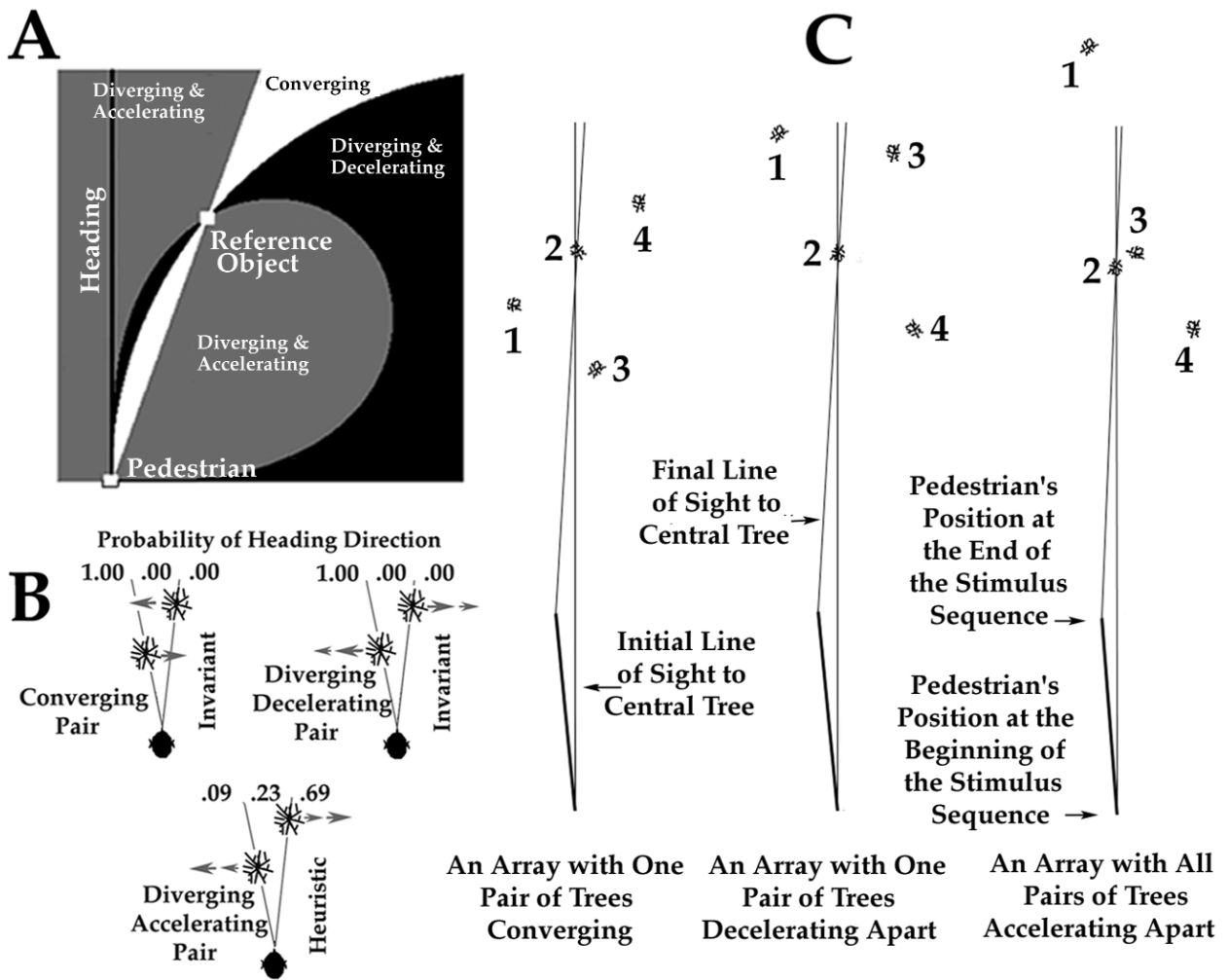


Figure 1. Plan views of a moving observer and surrounding stationary objects. Panel A shows a pedestrian and a given reference object off to the right of his or her path. Shaded areas denote the locations of all possible second objects and the types of motion they would have with the reference object. These motions occur regardless of whether or not the pedestrian fixates either. Panel B shows the statistical probabilities about heading direction associated with the three classes of motion, with arrows indicating direction of relative motion and diagonals indicating the lines of sight to the two trees demarcating three possible heading categories (to the outside of the near tree, between the trees, and to the outside of the far tree). Panel C shows three sample situations used in the experiment, with heading direction 8° to the left, always to the left of Tree 1. Trees are numbered left to right. The left panel shows a layout with Trees 3 and 4 converging throughout the course of the trial. All other pairs, adjacent (1–2, 2–3) and nonadjacent (1–3, 1–4, and 2–4), accelerate apart. Thus, the information specifies that heading direction is to the left of Tree 3. The middle panel shows a layout of a trial with Trees 2 and 3 decelerating apart; all others accelerate apart (1–2, 3–4, 1–3, 2–4, and 1–4). Heading direction is specified as being to the left of Tree 2. The right panel shows a situation in which all pairs (1–2, 2–3, 3–4, 1–3, 2–4, and 1–4) accelerate apart. In this case, heading direction is unspecified.

The third situation, on the other hand, is probabilistic: The existence of two objects accelerating apart offers no firm assessment of the location of one's heading. Wang and Cutting (1999) calculated that, in 9% of all such cases, one's heading direction is to the outside of the near object; in 23% of all cases, it lies between them; and in 69% of all cases, it lies outside of the far object. Here, any heading response to the outside of the far object would follow a *heuristic* (Gilden & Proffitt, 1989; Kahneman, Slovic, & Tversky, 1982)—a reasonable, but unsure, bet about where one's heading might lie. Cutting and Wang

(2000) analyze this situation further, but their results are beyond the scope of what is discussed here.

We investigated observers' eye movements during the course of their simulated travel toward a small grove of trees and prior to making a heading judgment. We anticipated that this behavior would show that observers seek out invariants. In particular, we made three predictions: (1) that observers would look more at trees within invariant pairs than at trees within heuristic pairs, because heading information is more trustworthy from the former than the latter, (2) that they would look less between the trees

of an invariant pair than between those in a heuristic pair, because heading can never lie between the former, and (3) that their gaze patterns would predict their heading responses and their relative accuracy.

METHOD

Stimuli

Motion sequences were generated at a median of 17 frames/sec on a Silicon Graphics Iris Workstation (Model 4D/35GT) with a UNIX-based operating system. The observers sat 0.5 m from the monitor, yielding 40°-wide displays at a resolution of about 30 pixels/deg. Perspective calculations were based on this vantage point. Sequences simulated 4 sec of linear translation of the observer over a flat plane at 1.25 eye heights/sec (about 2 m/sec for pedestrians 1.75 m in height, with their eyes 1.6 m above the ground) toward a grove of four trees. Arrays of four trees were chosen because they

offer several pairs of objects for the observer to consider but few enough to know where the observer is looking within the resolution of an eye tracker. Stimulus sequences began with the observer 14 eye heights from the middle of the grove and ended at about 9 eye heights. Trees in each stimulus sequence were 2.3 eye heights tall, with major branching at 1.1 eye heights. Each was identical in shape but rotated to a new random orientation for each tree on and each trial. The sky was light blue, the ground plane was brown, and the trees were black. The horizon was true, not truncated at a given depth plane. Four sample final frames are shown in Figure 2A.

All sequences kept one tree at midscreen during forward travel. This technique mimics the dolly and pan of a moving camera (Cutting et al., 1992; Kim, Growney, & Turvey, 1996; Royden, Crowell, & Banks, 1994). It also mimics the optical information during a pursuit-fixation eye movement, which predominates during pedestrian gait. Moreover, whether an observer gazed at the tree at mid-screen or not, there was always a dissociation between the feedback from eye muscles and what would occur viewing a natural scene

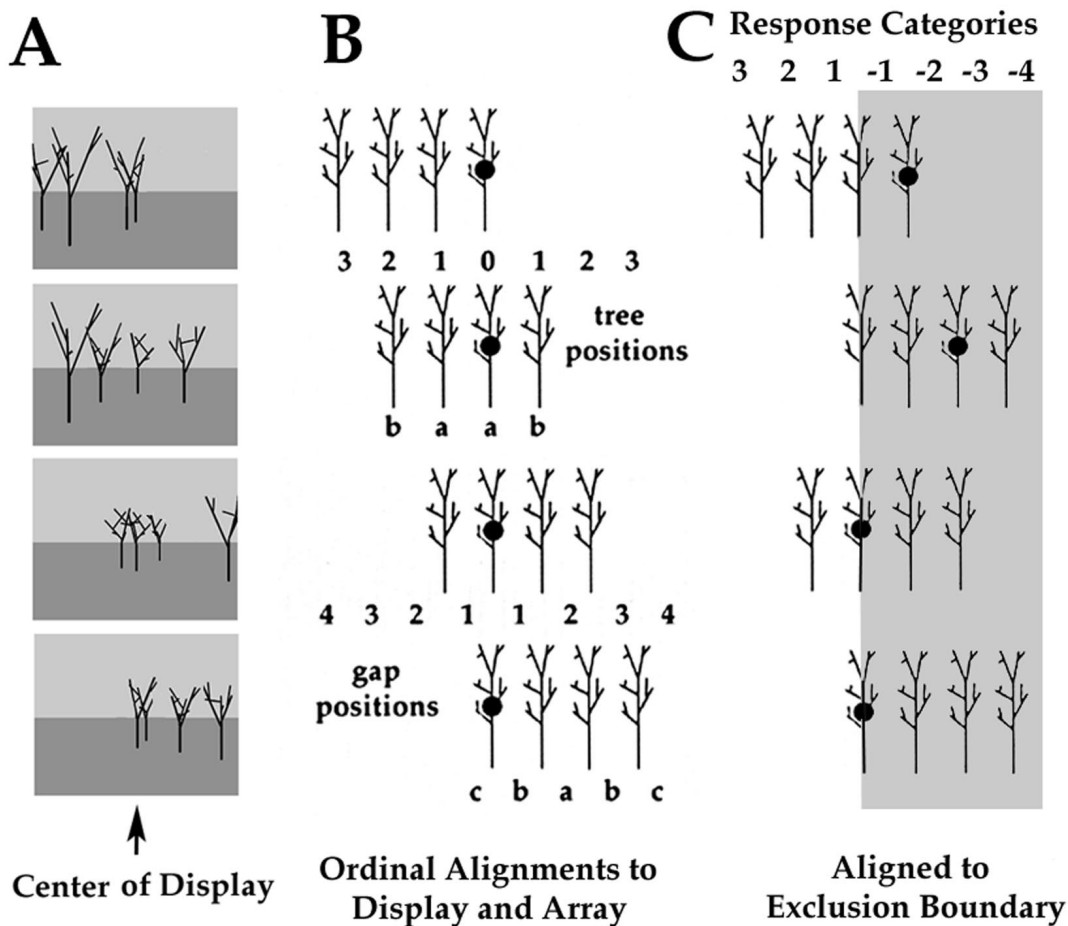


Figure 2. Actual and schematic arrangements of trees on given trials. Panel A shows sample final stimulus frames used in the experiment. Notice that different trees, counting left to right, are placed in the center of the display. Panel B shows the respective ordinal alignments around the central tree, shown with dots in their middle. Trees are denoted with numbers (0–3) counting outward from the center of the display and with letters (a and b) counting outward from the center of the four-tree array. Similarly, gaps are denoted with respect to the flanks of the tree at the center of the screen (1–4) and the center of stimulus array (a–c). Panel C shows a sample realignment of stimuli around exclusion boundaries, the near trees of invariant pairs (trees that converge or decelerate apart). The case shown in the third panel of Figure 2C is like that in the middle panel of Figure 1C. Such shifts are pertinent to the analysis of the distribution of heading responses in the gaps between trees (response categories), as in Figure 3A.

corresponding to the display. Such feedback has been shown to aid heading judgments when eye/head rotations are 2°/sec or more (Royden et al., 1994), but not when rotations are less, such as those used in the present study. Regardless, the purpose of this technique was to investigate optical information decoupled from systematic feedback from eye muscles.

Stimulus sequences were varied orthogonally in four ways. First, stimuli contained four different types of pairwise motions, created by varying the positions of the trees with respect to the observer's simulate heading. One stimulus set had at least one pair of trees converging, or moving toward one another, in the field of view. The left panel of Figure 1C shows a scaled, plan view of such a case in which Tree 3 and Tree 4, an adjacent pair, converge during the course of the trial. The relative motion of this pair denotes that the heading vector is to the left of Tree 3. A second trial type had at least one pair of trees diverging and decelerating away from each other. The middle panel of Figure 1C shows a case in which another adjacent pair, Trees 2 and 3, decelerate apart during the trial. Here, heading direction is denoted to the left of Tree 2. A third trial type had at least one pair converging and another pair diverging and decelerating. We call these three stimulus classes the *invariant stimuli*. A final set had all four trees diverging and accelerating away from each other. The right panel of Figure 1C shows the layout of such a trial. We call these *heuristic stimuli*. Over all trials, there were 122 trees involved in invariant pairs (either adjacent or nonadjacent) during some portion of the trial sequence and 134 trees that were not. These latter trees we designate as members of heuristic pairs.

Second, there were four arrangements of stimulus trees, each with a different tree (considering them right to left) in its initial position held at the center of the screen.¹ Examples are shown in Figure 2A. This experimental manipulation varied initial eye position within the stimulus array, allowing for saccades elsewhere, and partly overcame the general tendency for the observers to continue to look at midscreen. These stimuli are also plotted schematically in Figure 2B, in what we will call *ordinal arrays* (i.e., ignoring the absolute distances between trees). This normalization allows direct comparisons between various types of stimuli, with and without invariants, having controlled for the ordinal distribution of trees in the display. Figure 2B also shows the ordinal designations of the trees counting away from the center of the display (numbered 0 to 3) and the center of the array (a for the middle two trees and b for the trees on either end) and the designations of the gaps between trees counting out from the display center (numbered 1 to 4) and array center (a through c). Figure 2C aligns the trees in sample stimuli a different way for further analysis. In each case, a hypothetical near tree of an invariant pair is aligned, and response categories (gaps) are designated around it. These are numbered from -4 to +3, with +1 indicating the category just to the heading's side of the invariant pair. (The cases shown in the third panels of Figures 2B and 2C are like the trial shown in the middle panel of Figure 1C.) Positively numbered response categories are those allowed by the nominal invariants, negatively numbered categories are excluded, and the near tree of the invariant pair marks what we call the *exclusion boundary*.

Third, there were two initial heading eccentricities, 4° and 8° from midscreen. Corresponding final eccentricities were 6.2° and 12.4°. Thus, the maximum simulated eye/head rotation was 1.1°/sec, within the limits found by Royden et al. (1994) for accurate heading judgments without eye muscle feedback. Thus, feedback from real eye movements should offer no additional accuracy in determining heading. Fourth, trials simulated linear translation of the observer to the left or right of the tree at midscreen. Trials to the left were mirror images of those to the right.

Test sequences consisted of different random orders of the same 64 trials: 4 trial types (3 invariant sets + 1 heuristic set) × 4 arrangements of stimuli with different trees midscreen × 2 eccen-

tricitities × 2 approaches. Testing was preceded by 6 practice trials of sequences with 7 trees at an initial eccentricity of 8°. All such trials had at least one invariant pair of trees. After each practice trial, the observers were given nominal feedback, indicating whether a response was on the correct side of the tree at midscreen. No feedback was given during test sequences. At the end of each sequence, the last frame remained on the screen until the observer moved the mouse-controlled screen cursor to the place on the horizon toward which he or she thought he/she was headed and then clicked a mouse key. This response began the next trial.

Observers

Fifteen members of the Cornell University community participated individually for about 50 min and were offered \$10 for their services. Each viewed the 64-trial sequence twice, first without restraint and then with their eye movements and positions monitored by an Applied Science Laboratory Model 210 head-mounted limb eye tracker. Eye-movement records were noisy and uninterpretable for 3 observers. Of the remaining 12 observers, 2 were the first two authors, 1 was a laboratory assistant, and 9 were naive to the purposes of the experiment. All observers had normal or corrected-to-normal vision. To anticipate some results, there were no differences between naive and experienced observers.

Procedure

After viewing the first sequence, the observers had the eye tracker mounted, adjusted, and calibrated. Each sat with his or her head supported by a chinrest during the calibration and test. We then video-recorded test sequences at 60 Hz from the head-mounted camera, with cross hairs superimposed by the eye tracker, whose resolution was 1° measured horizontally. Each trial began with a 5-sec fixation cross at midscreen, during which the eye tracker could be recalibrated, if necessary. The observers were specifically instructed to look anywhere on the screen they liked during the course of the trial, to determine their heading during the course of the trial, and, at the end of the sequence, to move the screen cursor and press the mouse key to indicate their heading. The screen cursor appeared after the motion sequence ended.

Two types of response were recorded: heading judgments at the end of each trial and eye fixations throughout each trial. Videotape sequences of the observers' eye positions were scored off line. Scoring templates were created for each trial, indicating the positions of the four trees during every 100-msec interval. Stepping through the tape sequences frame by frame, we recorded horizontal eye position graphically. There was very little variation above or below the horizon. Scoring sheets were coded for dwell times and number of fixations on each tree and within each gap. Any gaze within 1° of the trunk of the tree in a given frame was scored as a fixation on the nearest tree; any gaze more than 0.5° off a tree was scored as a fixation within a gap. In this manner, some fixations were scored both ways.²

RESULTS AND DISCUSSION

We consider first the heading response data, then the eye movement data, and, finally, their conjunction. The heading responses were scored in two ways: absolutely (the screen position with respect to the center of the display) and categorically (in which of the five gaps the response was placed).

Heading Responses

Absolute heading judgments. We first assessed whether wearing the eye tracker altered heading re-

sponses, comparing the first and second runs through the sequences. We found no differences nor interactions involving conditions ($F_s < 1$). Mean response eccentricities were 2.2° and 2.4°, respectively, to the appropriate side of the fixation tree. Thus, there was nothing special about the observers' judgments while wearing the eye tracker. There was also no effect of side of approach, left or right ($F < 1$), so other analyses collapse across these data as well. As is common in this type of research, however, there was a reliable effect of eccentricity [$F(1,11) = 6.79$, $p < .03$] with mean cursor placements at 2.0° and 2.5° from the central tree, respectively, for initial 4° and 8° eccentricities. Since final eccentricities were 6.2° and 12.4°, these mean responses represent errors of 4.2° and 9.9°. Such results, among others, suggested to Cutting et al. (1999) that observers are not judging absolute heading in this stimulus context.

Most importantly, there was a reliable difference in the eccentricity of heading placements between invariant stimuli (2.6°) and heuristic stimuli (1.3°) [$F(1,11) = 20.6$, $p < .0001$]. This means that invariants guided the observers, in an absolute sense, to make more accurate responses. In addition, mean placements were 3.0°, 2.0°, and 2.9°, respectively, for invariant stimuli with converging pairs, diverging decelerating pairs, and both. Thus, in this analysis, convergence seems the more potent invariant.

Frey and Owen (1999) proposed a new measure of the information in a heading display, called the *separation ratio*. This ratio, σ , can be expressed as:

$$\sigma = 1 - (d_N/d_F), \quad (1)$$

where d_N is the distance from the observer to the near object at the end of the display sequence and d_F is the distance from the observer to the far object. Notice that σ always falls between 0.0 and 1.0, and the point of its measure is to predict that the relative amount of egocentric distance between the objects in an environment is the major factor in heading performance. Indeed, in our displays, the value of σ varied between .26 and .79, and was reliably correlated with mean eccentricities in heading judgments ($r = .41$) [$t(30) = 2.46$, $p < .02$]. However, in stepwise multiple regression with two predictors— σ and the presence (coded as 1) or absence (coded as 0) of invariants on any given trial—the invariant code accounted for 29% of the variance in the data ($r_i = .53$) [$t(30) = 3.46$, $p < .002$], and σ accounted for only an additional 6% ($r_{\sigma,i} = .27$) [$t(29) = 1.69$, $p > .10$]. Thus, the separation ratio is correlated ($r = .31$) with the invariant information in these displays but does not reliably contribute to the responses given in this context. The right and left panels of Figure 1C suggest why this might be the case. In the left panel, σ is .43 (where the near and far objects are Trees 3 and 4, respectively), and these two trees converge. In the right panel, σ is .48 (using Trees 4 and 1), yet all pairs of trees accelerate apart.

Categorical heading judgments. Consider the differences among the four types of trials, shifted in a man-

ner such as that shown in Figure 2C. For graphic simplicity, we again show correct heading judgments as if always to the left. For the invariant stimuli, all were shifted to the exclusion boundary, the location of the near tree of the invariant pair in the four-tree grove. Thus, in each stimulus, heading was specified as always left of this tree. Each invariant stimulus was matched with a heuristic stimulus that had the same ordinal alignment (central tree in the same ordinal position counting from left to right, as in Figure 2B), then the responses in the two groups were summed by response category (Figure 2C). Plots of these data are given in the panels of Figure 3A. More responses to the left of the boundary would be predicted in each case, and this occurred. There were reliable differences in each case: stimuli with converging pairs versus their heuristic counterparts [$F(1,11) = 64.6$, $p < .0001$], stimuli with a diverging decelerating pair versus their heuristic counterparts [$F(1,11) = 6.2$, $p < .03$], and stimuli with both converging and diverging decelerating pairs versus their counterparts [$F(1,11) = 28.4$, $p < .0001$]. Again, converging information seems more potent, but, in all cases, the invariant information shaped responses by leading the observers to place their heading estimates to the outside of the nearer tree.

Finally, we regressed the number of responses per gap across observers as a function of two variables: (1) whether or not the gap was included or excluded by the invariants (Figure 2C), and (2) the physical size of those gaps (measured in degrees of visual angle at the end of the trial). Across the 32 unique trials \times 5 gaps, there was a reliable effect of the invariants [$F(1,157) = 27.5$, $p < .001$] but not of physical gap size [$F(1,157) = 1.2$, $p > .25$]. This suggests our categorical analysis captures the essence of the data and that the angular distances between the trees is not a psychologically relevant variable.

Eye Movements, Fixations, and Dwell Times

Consider some overall characteristics of the eye recording data. First, the observers made an average of 6.75 fixations during the course of the 4-sec sequences (range across observers = 4.07 to 9.05). Mean dwell times were thus about 590 msec (range = 440 to 890 msec), considerably longer than is typical in reading or in tasks with static photographs or line drawings (Henderson & Hollingworth, 1998). Second, the observers spent most of their time (63%) looking at one of the four trees and considerably less time (24%) looking in gaps between trees. Given the resolution of the eye-tracking system, the residual dwell time (13%) could not be assigned definitively to either category and, to be conservative, was assigned to both in later calculations. Third, there were reliable effects of fixations and dwell times as a function of the egocentric depth of the trees in the array. The mean numbers of fixations were 1.28, 1.18, 1.08, and 0.93 for the nearest to the farthest tree, respectively; respective dwell times were 820, 742, 691, and 578 msec [$F_s(3,33) > 16.1$, $p_s < .0001$]. This latter result—that observers spend in-

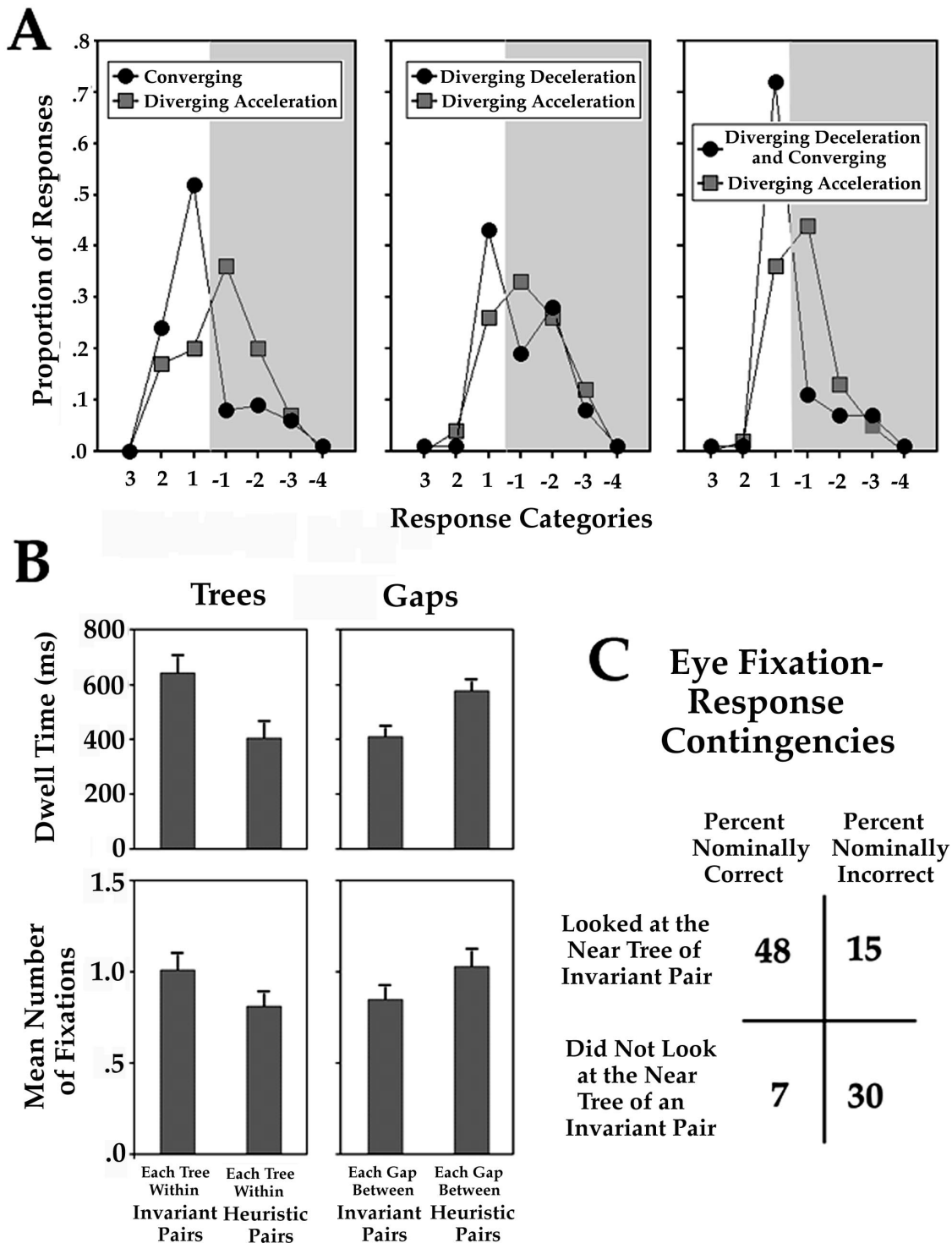


Figure 3. Responses, eye movements, and their conjoint results. Panel A shows the realignments of stimuli (as in Figure 2C), comparing each of the three types of invariant stimuli with their corresponding heuristic stimuli (those with the same arrangements of trees, as in Figure 2B). In each case, each invariant stimulus was aligned to the exclusion boundary and to its matched-pair heuristic stimulus. The panels of Figure 1C would be examples, where the right panel would serve as the heuristic match to the left panel for convergence and to the middle panel for decelerating divergence. Responses were then summed across stimuli for each response category (or gap). Panel B shows mean dwell times and mean numbers of fixations on each tree within invariant pairs and heuristic pairs and within each gap between invariant pairs and heuristic pairs. Error bars indicate 1 standard error of the mean. Panel C shows the proportions of responses to the correct and incorrect side of the near tree of invariant pairs as a function of whether the observers looked at that tree during the course of the stimulus sequence.

Table 1A
The Numbers of Trees and Gaps in Various Combinations
Across the 48 Invariant Stimuli, Mean Dwell Times
(in Milliseconds), and Number of Fixations

Display Position	Array Position	Members of Invariant Pairs			Member of Heuristic Pairs		
		<i>N</i>	Dwell	Fixation	<i>N</i>	Dwell	Fixation
0	a	22	1,600	2.31	2	1,320	1.96
	b	13	1,390	2.16	11	1,130	1.98
1	a	43	930	1.51	13	900	1.45
	b	11	370	0.50	13	150	0.24
2	a	17	490	0.80	7	280	0.58
	b	7	170	0.36	17	190	0.33
3	b	9	190	0.36	15	60	0.12

Table 1B
The Number of Gaps Between Trees in Various Combinations
Across 48 Invariant Stimuli, Mean Dwell Times
(in Milliseconds), and Number of Fixations

Display Position	Array Position	Between Invariant Pairs			Between Heuristic Pairs or Outside the Array		
		<i>N</i>	Dwell	Fixation	<i>N</i>	Dwell	Fixation
1	a	18	580	1.13	3	830	1.65
	b	25	510	1.13	29	800	1.36
	c	—	—	—	23	250	0.36
2	a	17	290	0.62	6	500	1.03
	b	10	300	0.63	11	410	0.63
	c	—	—	—	21	40	0.05
3	b	11	180	0.42	12	260	0.43
	c	—	—	—	21	60	0.07
4	c	—	—	—	23	30	0.06

Note—Display and array positions refer to those in Figure 2B.

crementally less time looking at farther objects—was also found by Wagner et al. (1981) for gaze patterns of pedestrians walking through a campus and college town.

Consider next eye movements as a function of stimulus variables, first concerning tree fixations, then fixations within the gaps between trees. Here, we used only those 48 trials with invariants present, because we wanted to compare the dwell times and fixations within these trials. We carried out all analyses separately on individuals, pooling results for stimuli within each case, then tested differences across individuals. In this manner we factored out any differences between the ordinal locations of trees within invariant versus heuristic pairs. The number of stimuli with trees or gaps in each case for each individual is given in Tables 1A and 1B.

Tree fixations. Again, tree positions were numbered with respect to their ordinal distance from midscreen (0 through 3, as shown in the upper panel of Figure 2B). The observers spent incrementally more time looking at trees toward the center of the display than farther away [$F(3,33) = 62, p < .0001$]. Mean dwell times were 1,420, 700, 310, and 110 msec, respectively, for trees at locations 0 through 3. There were also more fixations nearer midscreen [$F(3,33) = 231, p < .0001$], with means of 2.18, 1.12, 0.54, and 0.21 fixations per tree in each lo-

cation. Neither result is surprising since we instructed the observers to fixate the middle of the screen before each trial began. Inertia would keep them there. In addition, trees in the middle of the array (Trees a in second panel of Figure 2B) were glanced at longer and more often than trees at the ends of the array (Trees b): 960 msec and 1.50 fixations on each Tree a, and 460 msec and 0.75 fixations on each Tree b [$F_s(1,11) > 14.8, p < .003$]. This, too, was of little surprise: Indeed, 85% of all first saccades were toward the middle of the stimulus array (range = 75% to 95%).

As we predicted, the observers looked 240 msec longer at [$F(1,11) = 14.8, p < .003$] and produced 0.18 more fixations on [$F(1,11) = 28.5, p < .0001$] each tree that was a member of invariant pairs than on each other tree, as shown in Figure 3B.³ Regression analyses revealed no differences among the three classes of invariant stimuli. Separate means for trees in different display positions, array positions, and associated with invariants or heuristics are shown in Table 1A. There were also no differences in the number of invariant versus heuristic trees at different depths.

Gap fixations. Consider again the lower panels of Figure 2B. Gaps are denoted 1–4, away from the flanks of the tree at the center of the screen. Not surprisingly, dwell times and fixations within gaps decreased with distance from the screen's center. Mean durations were 580, 310, 120, and 40 msec; mean numbers of fixations were 1.06, 0.50, 0.22, and 0.06 [$F_s(3,33) > 68, p_s < .001$]. In addition, dwell times and fixations declined with distance from the center of the array (Gaps a–c); mean durations were 520, 480, and 160 msec; and mean fixations were 1.04, 0.78, and 0.11 [$F_s(2,22) > 17, p_s < .001$].

Consider next the eye position data in gaps between invariant pairs of trees versus those between others. As we predicted, mean dwell times were 170 msec shorter for gaps between invariants pairs [$F(1,11) = 11.7, p < .006$], and there was a mean of 0.18 fewer fixations there as well [$F(1,11) = 4.4, p < .06$], as shown in Figure 3B. Note that the inclusion of some fixations in both tree and gap analyses would only serve to dilute the differences shown in Figure 3B. Finally, across all stimuli and all gaps, physical gap size (measured in degrees of visual angle at the end of each trial) was negatively correlated with mean dwell times and mean number of fixations ($r_s < -.26, N_s \geq 192, p_s < .0001$). This was true regardless of whether one included the generally wider gaps on either end of the arrays (Gaps c in Figure 2B) or considered only the three central gaps (Gaps a and b). Thus, eye positions were clearly not distributed uniformly across the display but were concentrated on and near the trees.

Eye Movements and Heading Responses

One way to discern whether the observers sought their heading is to consider dwell times and fixations in the gaps within which the observers later placed their heading responses versus the other four in each array. We found

that the observers spent considerably more time looking within ($M = 730$ msec) and glanced more toward ($M = 1.25$ fixations) their response gap than the means of the other four gaps (250 msec and 0.48 fixations) [$F_s(1,11) > 21.6$, $ps < .001$]. Thus, it is clear that the observers frequently looked toward their judged heading before they responded with the mouse-controlled screen cursor.

Most important, however, is the relation between dwell times and fixations on the information-carrying trees prior to their heading response. Of the 48 trials with invariants, 20 of them included the tree in the center of the display as the near member of the invariant. Since the observers were told to begin each trial fixating this tree, we excluded these trials. For the remaining 28 trials, we tallied a 2×2 table for each observer: whether they looked at the near tree of the invariant pair or not, and whether they placed their heading response to the correct side of this tree or not. Overall percentages or responses are shown in Figure 3C. When the observers looked at the pertinent tree, they were correct 76% of the time in this sample; when they did not look at this tree, they were only 19% correct. Eight of the 12 observers showed reliable such patterns [$\chi^2_s(1) > 3.84$, $ps < .05$], and the other 4 observers also favored the major diagonal.

CONCLUSIONS

Our approach is different than others in the literature. One major difference is that most others focus on optical motion alone, irrespective of the particular objects that are moving. Ultimately, these theories appeal to the motion-sensitive receptive fields of cortical cells (e.g., Hildreth, 1992; Perrone & Stone, 1994; Rieger & Lawton, 1985). Some, following Gibson (1979), look specifically for the efficacy of radial expansion patterns (Duffy & Wurtz, 1991; Tanaka, 1998), sometimes in the context of other motions (Bradley et al., 1996), but few consider motion in conjunction with depth (but see Frey & Owen, 1999, and Perrone & Stone, 1994), and none have considered the motions of attended versus unattended objects. Since pedestrian speeds are those to which we adapted, and since these yield quite modest global optical motions, radial expansion patterns are typically not salient except when driving a sports car or landing an airplane. Thus, pedestrians must actively seek out information about heading; it is not a passively received fact of locomotion. In seeking out heading information under such conditions, attention and eye movements are needed. The attention paid to where, and near where, one looks is likely to suppress the motion responses of many cells and enhance those of others (Motter, 1993; Mountcastle, Motter, Steinmetz, & Sestokas, 1987).

A pedestrian's search for reliable information about their heading manifests itself in four ways in our experiment. First, our data show that they look more often and longer at objects that belong to pairs that converge or decelerate

apart than at objects belonging only to pairs that accelerate apart. The rationale for this behavior seems apparent: As shown in Figure 1B, converging and diverging decelerating pairs offer trustworthy information that specifies heading direction, whereas diverging accelerating pairs do not. Second, pedestrians look less often in the gaps between pairs that converge or decelerate apart. The rationale for this is also apparent: As equally shown in Figure 1B, one's heading can never lie between such pairs. Third, heading responses appear to be guided by the nominal invariants in the stimuli. Distributional analyses, shown in Figure 3A, demonstrate that more responses were placed to the outside of the near tree of an invariant pair than outside the corresponding trees in heuristic pairs. Finally, accurate responses were contingent on the observers' having looked at invariant pairs. On trials selected so that the observers did not begin by fixating on members of invariants pair, as shown in Figure 3C, performance faltered; the observers performed four times better when they looked at the near member of an invariant pair than when they did not.

To be sure, performance is not perfect in the presence of invariants. We know of no study in any domain in perception in which it is. Invariants specify what is to be perceived by an alert and attending observer. If the motions that comprise them are below some threshold, if the observer has simply not been lucky enough to discern their presence under the pressures of the task, or if the observer has not learned to differentiate them (as might be the case for the very young child), then performance will falter. That observers can find them and use them in this task without feedback supports both the naturalness of our study and the plausibility of our account. That observers are systematically wrong when they do not find them further supports our ideas. It is always possible that other schemes based on other sources of information can account for heading judgments in other situations, but Cutting et al. (1999) demonstrated that many of the schemes cannot account for heading responses in situations such as those studied here. Moreover, those other accounts offer no account of eye movements prior to heading judgments.

REFERENCES

- Bradley, D. C., Maxwell, M., Andersen, R. A., Banks, M. S., & Shenoy, K. V. (1996). Mechanisms of heading perception in primate visual cortex. *Science*, *273*, 1544-1547.
- Cutting, J. E. (1996). Wayfinding from multiple sources of local information in retinal flow. *Journal of Experimental Psychology: Human Perception & Performance*, *22*, 1299-1313.
- Cutting, J. E., Springer, K., Braren, P. A., & Johnson, S. H. (1992). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: General*, *121*, 41-72 & 129.
- Cutting, J. E., & Wang, R. F. (2000). Heading judgments in minimal environments: The value of a heuristic when invariants are rare. *Perception & Psychophysics*, *62*, 1146-1159.
- Cutting, J. E., Wang, R. F., Flückiger, M., & Baumberger, B. (1999). Human heading judgments and object-based motion information. *Vision Research*, *39*, 1079-1105.

- Duffy, C. J., & Wurtz, R. H. (1991). Sensitivity of MST neurons to optic flow stimuli: I. A continuum of response selectively to large-field stimuli. *Journal of Neurophysiology*, **65**, 1329-1345.
- Frey, B. F., & Owen, D. H. (1999). The utility of motion parallax information for the perception and control of heading. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 445-460.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gilden, D. L., & Proffitt, D. R. (1989). Understanding collision dynamics. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 372-383.
- Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269-293). Amsterdam: Elsevier.
- Hildreth, E. C. (1992). Recovering heading for visually-guided navigation. *Vision Research*, **32**, 1177-1192.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Kim, N.-G., Growney, R., & Turvey, M. T. (1996). Optical flow not retinal flow is the basis of wayfinding by foot. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1177-1192.
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, **70**, 909-919.
- Mountcastle, V. B., Motter, B. C., Steinmetz, M. A., & Sestokas, A. K. (1987). Common and differential effects of attentive fixation upon the excitability of the light-sensitive neurons of the parietal-prestriate (V4) cortical visual neurons in the macaque monkey. *Journal of Neuroscience*, **7**, 2239-2255.
- Perrone, J., & Stone, L. (1994). A model of self-motion estimation within primate visual cortex. *Vision Research*, **34**, 1917-1938.
- Rieger, J. H., & Lawton, D. T. (1985). Processing differential image motion. *Journal of the Optical Society of America A*, **2**, 354-360.
- Royden, C. S., Crowell, J. A., & Banks, M. S. (1994). Estimating heading during eye movements. *Vision Research*, **34**, 3197-3214.
- Tanaka, K. (1998). Representation of visual motion in the extrastriate visual cortex. In T. Watanabe (Ed.), *High-level motion processing* (pp. 295-313). Cambridge, MA: MIT Press.
- Wagner, M., Baird, J. C., & Barbaresi, W. (1981). The locus of environmental attention. *Journal of Environmental Psychology*, **1**, 195-201.
- Wang, R. F., & Cutting, J. E. (1999). Where we go with a little good information. *Psychological Science*, **10**, 71-75.
- Warren, W. H. (1995). Self-motion: Visual perception and visual control. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (pp. 263-325). San Diego: Academic Press.
- Warren, W. H. (1998). The state of flow. In T. Watanabe (Ed.), *High-level motion processing* (pp. 315-358). Cambridge, MA: MIT Press.

NOTES

1. Of the 64 stimulus sequences, 10 had two trees exchange ordinal positions (counting from left to right) after a period of convergence. When this happened, the invariant pair became one that accelerated apart and thus became a heuristic pair. Analyses treated such pairs as invariant pairs before crossover and heuristic pairs after. Cutting and Wang (2000) analyze this situation in further detail.
2. The different criterion for gaps was used because, during the course of many trials, many tree pairs were closer than 2°. Without a standard of less than 1°, any fixation between such pairs would automatically be recorded as a tree fixation. A standard of 0.5° allowed us to consider possible gazes between such pairs, which were quite prevalent.
3. Notice that this effect cannot be due to the fact that there were more trees in invariant pairs (122) than those in heuristic pairs (70) in these 48 stimuli. First, the data in Table 1A show longer dwell times and more fixations for matched samples of invariant and heuristic pairs at six of seven display position/array position combinations. Second, the analyses were done on the means for all trees of each type, not on total dwell times or fixations per trial. Thus, the results are corrected for the differing numbers of trees.

(Manuscript received October 28, 1998;
revision accepted for publication September 14, 1999.)