Task demands and binocular eye movements

Andrew E. Welchman
School of Biology (Psychology),
University of Newcastle-upon-Tyne, UK, and
Max-Planck-Institut für biologische Kybernetik,
Tübingen, Germany

Julie M. Harris
School of Biology (Psychology),
University of Newcastle-upon-Tyne, UK

Humans make rapid movements of their eyes several times a second that enable them to examine objects located at different positions in space with both of their eyes. Much of our understanding of these binocular movements comes from studies using experienced observers performing repetitive, unnatural tasks. But what eye movements are made when naïve observers perform tasks demanding specific binocular visual information? We examined the binocular eye movements produced by observers performing two tasks differing in the visual information needed for their completion. Our motivation for doing this was to examine the role and function of binocular eye movements when making decisions. We considered the fixation strategies adopted by observers, the effects of the task on the dynamics of saccadic eye movements, and the combination of vergence and version in gaze shifts. We report that the task-dependent use of visual information can have a strong influence on the patterns of fixations, whilst not influencing saccade dynamics. Our data provide some support for the notion that observers choose and fixate a notional reference point in the scene when making judgments about depth structure.

Keywords: binocular eye movement, vergence, disparity scaling, distance judgment

Introduction

Most human activity is accompanied by the repositioning of the fovea several times every second. The rapid eye movements that achieve this allow observers to sample the visual properties of objects located at different locations. But what specifically are such eye movements for? Do they simply help sample the world, or do their speed and target location, and any differences in movements between the two eyes, aid the brain in the processing of visual information? How does the task an observer is performing affect patterns of eye movement? Here we address these questions by examining the eye movements produced by naïve observers performing two different tasks where the same visual information is put to different uses. Before making predictions about the types of eye movement that might be expected from our chosen tasks, we briefly consider the background to our knowledge of binocular eye movements.

The Ecology of Eye Movements

For optimal binocular visual inspection of objects, observers need to change the vergence and version state of their eyes. Much of what we know about the nature of binocular eye movements has come from studies using simple stimuli (pin heads or LEDs) with highly constrained task demands (e.g., asking observers to look back and forth between visual stimuli at regular time intervals). This approach simplifies the study of eye movement. But it is based on the premise that eye movements recorded under such experimental constraints reflect natural eye movement behavior. This assumption is not always justified (for discussion, see Steinman, Kowler & Collewijn, 1990; Collewijn, Steinman, Erkelens, Pizlo & van der Steen, 1992). During the past decade, a number of investigators have attempted to establish (or revive Yarbus’ attempt at) an ecological approach to the study of eye movements. These studies have not been concerned with ocular-motor dynamics, but rather the role that eye movements play in controlling the selective sampling of visual information during complex tasks. These studies have focused on the integration of eye movements in complex sequences of activity: for instance, whilst driving a car (Land, 1992; Land & Lee, 1994; Land & Tatler, 2001), playing sports (Land & Furneaux, 1997; Land & McLeod, 2000), copying complex colored objects (Ballard, Hayhoe, & Pelz, 1995; Hayhoe, Bensinger, & Ballard, 1998; Pelz, Hayhoe, & Loeb, 2001) or performing everyday activities, such as making a cup of tea (Land, Mennie, & Rusted, 1999), making a sandwich (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003), washing one’s hands (Pelz & Canosa, 2001) or navigating a corridor (Turan, Geruschat, & Baker, 2003). These studies all emphasize the role of eye movements in allowing the brain to obtain information necessary for the cognitive task at hand. For instance, the study of Land et al. (1999) on the complex action sequence of making tea points out the importance of re-fixations on task-related objects for the online...
control of tasks requiring little cognitive effort. Although observers appeared not to be explicitly aware of it, their eyes were carefully repositioned to locate objects that they were going to manipulate, to direct movements of their limbs toward the objects, and to provide feedback on their actions.

One of the distinct advantages of this more naturalistic approach to the study of eye movements is that observers become engaged in the task they are performing and are thus more likely to make eye movements typical of the ones they make in everyday circumstances. Smeets, Hayhoe, and Ballard (1996) make this point explicitly by suggesting that eye movement experiments conducted under strict experimental conditions can be more informative about the constraints placed on the observer than on natural eye movements. However, a disadvantage of the naturalistic approach comes from the technical limitations of recording eye movements in freely moving observers: temporal and spatial sampling rates tend to be limited and recordings are rarely binocular. This ecological approach to the study of eye movements has told us much about the function of eye movements in everyday tasks, but little about the ocular-motor dynamics of daily life.

**Task-Specific Effects on Eye Movements**

One experiment in the literature marks a halfway point between the two approaches to the study of eye movements described above. Epelboim et al. (1995) used a unique apparatus to measure eye and head movements in (relatively) unconstrained subjects performing two simple tasks. The apparatus, the Maryland revolving field monitor (MRFM) (for a full description see Epelboim et al., 1995), enabled them to measure eye-in-head position, head rotation, and head-in-space position with great accuracy (1 arcmin for eye and head rotations, and 1mm for head position). The experiment they conducted was relatively simple: they asked subjects either to look at a series of LEDs positioned on a workbench in front of them (LOOK task) or to reach out and tap the LEDs in turn (TAP task). A number of different aspects of the data have been presented (Epelboim et al., 1995; Epelboim, 1998; Malinov, Epelboim, Herst, & Steinman, 2000; Herst, Epelboim, & Steinman, 2001), perhaps the most surprising of which was that gaze-shift dynamics were considerably different in the two tasks (Epelboim et al., 1997). Specifically, Epelboim et al. (1997) reported that in the TAP task, peak gaze-shift (eye and head movement) speeds and peak eye movement speeds (eye in head movement) were higher, whilst the duration of gaze-shifts was shorter in the TAP task than in the LOOK task. This is particularly surprising because saccadic eye movements have stereotyped, machinelike dynamics that are not under conscious control. Epelboim et al. (1997) suggested that task demands placed on the subject might influence saccade dynamics.

In this study, we further examine the role that task demands play in the production and control of binocular eye movements. In particular, we consider the role that vergence eye movements play when observers perform different tasks. We were particularly interested in this issue because vergence state potentially provides an important source of extra-retinal information that could be used to interpret retinal information about the three-dimensional (3D) location of objects in peri-personal space (Foley, 1980). We set up two tasks: a DEPTH task where subjects make a judgment about the 3D arrangement of three patches, and a RATIO task where observers made a judgment about the aspect ratio of the patches presented to them. The DEPTH task in principle could benefit from vergence information, whereas the RATIO task would not. Our aim was to test whether different patterns of eye movements would be obtained when observers performed the two tasks. Below we outline how vergence eye movement information could be useful.

**Eye Movement Strategies for Depth Judgments**

There are many situations in which retinal information by itself provides insufficient information about the layout of objects in 3D space. For example, the perception of 3D shape from horizontal binocular disparity information requires an estimate of viewing distance. This estimate of viewing distance is needed because depth (Z) is related to disparity (δ) by the following approximate equation (Howard & Rogers, 2002):

\[
\delta \approx \frac{IZ}{D^2},
\]

where I is the interocular separation, and D the viewing distance. Notice that if D is changed (e.g., by moving away from an object), the same depth, Z, will be specified by a different binocular disparity. How does the visual system obtain an estimate of viewing distance that can allow it to scale disparity for depth and shape perception? The vergence state of the eyes is potentially useful as there is a direct relationship between vergence angle and object distance. If observers had knowledge about the vergence posture of their eyes, estimates of object distance could be made from 10 cm to 6 m (Foley, 1980). Some evidence in the literature has cast doubt on the utility of vergence information as a cue to distance (Erkelens & Collewijn, 1985; Collewijn & Erkelens, 1990). However, extra-retinal cues to distance are no noisier than extra-retinal cues to visual direction (Brenner & Smeets, 2000), extra-retinal cues appear to contribute to distance perception (Brenner, van den Berg & van Damme, 1996), and some
evidence suggests that extra-retinal cues can support reliable judgments of distance (Mon-Williams & Tresilian, 1999).

The two tasks used in this study differed in the necessity of subjects to estimate viewing distance for their successful completion. Our first task was devised such that it required observers to make an estimate of viewing distance. This DEPTH task involved observers making a judgment about the shape of three patches positioned in depth to form an isosceles triangle (see Figure 1). The apex of the triangle was closer to the observers than the base, and observers were required to indicate whether the base of the triangle was longer than its height (apex to base distance). This task was modeled after the Apparently Circular Cylinder task and its variants used elsewhere to investigate the use of depth cues – Johnston, 1991; Glennerster, Rogers & Bradshaw, 1996; Bradshaw, Parton & Glennerster, 2000). This task requires that observers make judgments about the metric depth structure of the information presented to them.

There are potentially three ways in which observers could exploit vergence eye movements to obtain information about the distance of objects in a scene. First, Enright (1996) noted that observers judging the relative distances of objects tend to look at them in turn. One way to obtain information about the layout of objects in a scene would be to lock the vergence state of the eyes across a saccade so that when the eyes landed on a target at a different distance, the object would fall at disparate points on the retina. The difference between the retinal position of the object following a saccade (absolute disparity with respect to the fovea), compared with where it was before the eye movement, could be used to measure disparities (Enright, 1991). This could be advantageous as very large disparities of objects far apart in both depth and lateral distance could be registered with respect to the fovea rather than being measured as much more peripheral relative disparities. This would also be useful if the extra-retinal vergence information was not accurate or precise. It could be easier to monitor maintained vergence than to measure shifts of vergence. Under this scheme, observers would be expected to make isovergent saccades when making distance judgments.

Second, Foley (1980, 1985) suggested that the visual system uses a single estimate of distance, a notional reference point, relative to which differences in binocular disparity are scaled to obtain estimates of depth. It is not known whether observers specifically fixate on their chosen reference point. If observers were to use such a scheme they might be expected to fixate that reference point, obtain an estimate of the distance of the point, perhaps through sustained vergence, and then make judgments about the position of other objects with respect to this fixated point. It would not be necessary to move the eyes again.

Third, observers often combine changes of version with changes of vergence (e.g., Ono, Nakamizo & Steinbach, 1978; Enright, 1984, 1986, 1992; Collewijn, Erkelens, & Steinman, 1988a, 1988b). Observers may be sensitive to changes in ocular convergence across saccades (Wright, 1951; Brenner & van Damme, 1998). If this is the case, observers making judgments requiring distance scaling could simply make saccades that coordinate changes of version and vergence and register the changing vergence component. Under this scheme, observers making distance judgments would be predicted to make regular saccades that combine vergence changes with shifts in version.

The second task we employed was one that would not benefit from accurate vergence changes. In the RATIO task, observers examined three rectangular patches and determined which of them had a different aspect ratio from the other two presented (see Figure 1). This task can be performed on the basis of retinal information:

**Stimulus configuration:**

- **Ratio task**
  - **Reference:**
  - Which patch has a different aspect ratio?
  - Left or right?

- **Depth task**
  - **Which dimension of the triangle is longer?**
  - Z or X?

![Figure 1. A description of the tasks. Observers were presented with rectangular patches positioned in depth to form an isosceles triangle with the apex nearest to them. They were required to perform different tasks on the same visual information. The RATIO task involved making a judgment about the aspect ratio of the luminous patches they saw. Using the central patch as a reference, they had to decide whether the left or the right patch had a different aspect ratio (i.e., the ratio of the patch’s width to height). The DEPTH task involved observers making a judgment about the triangular configuration of the patches in depth. They were required to determine whether the triangle was longer in the z dimension (apex to base distance) or in the x dimension (base width).](jov.arvojournals.org)
observers could simply compare the retinal width of a patch to its retinal height. In this task, the distance of each rectangle from the observer is functionally irrelevant. To be precise, it should be noted that although the physical aspect ratios of the targets might be equal, on the retina they would be slightly distorted due to perspective. We calculated the perspective distortion of the aspect ratios for our task configuration and found them to be very small indeed (maximum distortion was 0.8%). Such distortions would be indiscriminable as they are well below the aspect ratio discrimination thresholds (Regan & Hamstra, 1994: thresholds are around 3-5%). Therefore, we are confident that all the useful information for performing this task is specified on the retina. For this task, we expected observers to fixate sequentially on each of the patches to examine the aspect ratio of each patch with their fovea. Control of vergence state across saccades would need to be sufficient only to prevent diplopia.

Using the two tasks described above, the aims of the study were, first, to characterize the use of eye movement strategies produced by naïve observers when performing a task requiring judgments about shape involving the scaling of disparities; second, to compare these measurements to those obtained when subjects performed a task that did not require accurate binocular eye movements, to determine whether there were any task specific effects (cf. Epelboim et al., 1997; Malinov et al., 2000); and, finally, to confirm that eye movement dynamics measured when observers perform a task unrelated to eye movement are similar to those when subjects are given specific instructions about the movement of their eyes.

**Methods**

**Stimuli**

Stimuli consisted of rectangular patches of phosphorescent paint glued to wooden mounts. The patches were created by printing rectangles of known size onto self-adhesive envelope labels. The rectangles were then painted using an acrylic, phosphorescent paint (“luminous glow in the dark” paint; Inscribe, UK), and then carefully cut out using a craft knife. The phosphorescent patches were stuck onto wooden mounts (4 by 4 cm square) so they were equidistant from the left and right edges. The vertical position of the patches on the mounts was pseudo-randomized (within the 4-cm height of the mount) to remove the height-in-the-field cue to distance. The patches were “charged” under a 60-W desk lamp for 5 min prior to use.

Sixteen different rectangular patch stimuli were used in the experiments. Patches varied in height (3, 2.5, 2, or 1 cm) and also in the ratio of width-to-height (1:1.5, 1:1.8, 1:1.9, 1:2, 1:2.1, 1:2.2, and 1:2.5). Three rectangular patches were presented on any given trial: the patch in the middle (the reference) always had an aspect ratio of 1:2; one of the other patches also had an aspect ratio of 1:2, and the other patch had a different aspect ratio. The three presented patches never had the same height, so it was not simply possible to make a discrimination based on the width of the object. The three patches were configured to form an isosceles triangle that lay in the transverse plane of the head (see Figure 2). The apex of the triangle was 50 cm from the observer in the mid-sagittal plane. The depth of the triangular configuration (base to apex distance) was varied between 12 and 16 cm; the width of the base was independently varied between 12 and 16 cm (see Figure 2). This manipulation approximately corresponded to horizontal disparities of 1.47 to 1.85 deg between the nearest and farthest points of the configuration.

**Apparatus**

All stimuli were presented within a viewing box that is depicted in Figure 2. The box had adjustable baffles 35 cm from the observer’s eyes that could be moved to ensure that the observer could not see the sides or the ceiling of the box. Stimuli were manually placed in the box via the access portals located on the side of the box.

![Figure 2. A cartoon of the experimental set up.](Downloaded from jov.arvojournals.org on 03/17/2019)
The interior of the box was painted matte black. The floor of the box was marked with a square grid drawn using an ultraviolet sensitive security marker pen. Using an ultraviolet fluorescent lantern, the experimenter could precisely position stimuli between trials; during a trial the positioning grid was totally invisible to the observer. The viewing box was equipped with a shutter device 30 cm from the observer’s eyes. The gravity-driven shutter consisted of a pair of mounted solenoids attached to a ridged bar that acted as the mount for a matte black fabric curtain. Demagnetizing the solenoids caused the curtain to fall, enabling the observer to see the stimuli (fall duration ca. 300 ms).

Eye movement recordings were made using an infrared video-based binocular eye tracker (Applied Science Laboratories [ASL] binocular model 501 with high speed cameras). The tracker works by detecting and tracking the centroid of the observer’s pupil and the first-order corneal reflection of the infrared light source. Eye position was sampled at 120 Hz in each eye. The ASL 501 system has a quoted positional accuracy of 0.5 deg and a positional resolution of 0.1 deg. The eye tracker was desk-mounted; the subject’s head was stabilized using a chin rest and forehead restraints. The tracker was calibrated for both eyes using a 3 x 3 grid of points located 58 cm from the observer’s eyes. The calibration area was 16 deg horizontal and 5.5 deg vertical. Eye position data were recorded by a PC interfaced with the eye tracker’s control modules through two serial port connections. Eye position data were stored on the PC’s hard disc; analysis was performed off-line (described below). The ASL system also includes a scene camera mounted at approximately the bridge of the observer’s nose. The system superimposes a crosshair on the scene image at the observer’s position of gaze. This scene image was used to check calibration, and we also made a recording of each trial using a video recorder interfaced with the eye tracker. However, eye position data, rather than scene images, were used for data analysis.

Tasks

Observers were required to perform two different tasks. The RATIO task involved comparing the aspect ratios of the phosphorescent patches. Observers were told that the reference patch was always presented in the center, and they had to indicate (by knocking on the desk with their hand) whether the patch on the left or the right had a different aspect ratio from the reference. The DEPTH task involved observers judging the triangular configuration of the three patches. They were required to indicate (by knocking) whether the triangle was longer in depth (z distance in Figure 1) than in width (x distance in Figure 1). Observers performed two blocks of 16 trials; during a block, the task they performed (RATIO or DEPTH) was the same. The order of blocks of trials was randomized between observers.

Procedure

Observers were given an instruction sheet with a schematic diagram informing them about the task they were about to perform. They were not shown any stimuli, but were questioned to ensure they had grasped task requirements.

Observers were positioned in front of the viewing box so their eyes were level with the box’s floor. The head was secured using a chin rest and restraints on the observer’s temples. The calibration card was lowered into the viewing box through a slot in the roof (see Figure 2) and the eye tracker calibrated. Illumination for calibration was provided by 40-W light bulbs located behind the baffles on each side of the box. The shutter was set, the calibration card removed, and stimuli placed in the viewing box.

A trial consisted of the experimenter dropping the mechanical shutter, starting eye position recording, and waiting until the observer made a response before stopping data acquisition. The subject’s response was recorded, after which the shutter was reset, and new stimuli were placed in the viewing box.

Observers performed 16 trials, after which eye tracker calibration was checked, and they were given a brief rest. The next task was explained to the observer; they were then repositioned in the eye tracker and recalibrated. As we were interested in the possibility that the production of eye movements involves a degree of learning about the structure of the scene, exactly the same stimuli were shown to the observers in the two blocks of the experiment in the same presentation order. (Learning effects are not addressed further because we found no evidence for any such effects). Subjects were aware that their eye movements were being recorded; however, the instructions they were given related only to the task, and eye movements were not mentioned.

Eye Movement Analysis

Eye position records were analyzed off-line using custom-written software. Analysis software recalibrated eye position recordings made by ASL’s eye tracker control software using the calibration data obtained at the start of an experimental block (the method involved linear interpolation in the horizontal and vertical directions). Vergence angles were calculated as the difference in the horizontal angular position of the two eyes added to the vergence angle of the eyes at the distance of the calibration card. Our aims in the eye movement analysis were (i) to remove blinks, (ii) to identify saccades, (iii) to classify fixations, and (iv) to register changes of vergence.

Blinks in the data were detected by noting a pupil diameter of zero in either eye. Eye position signals were unreliable following the occurrence of a blink, for this reason 75 ms of eye position data were removed following a blink. After removing blinks from the data, the first-
order temporal derivative of the horizontal and vertical positions of each eye and the first-order temporal derivative of vergence angle were calculated by differentiating using a numerical method (calculating the slope of the eye position trace between the two adjacent points for each data point). Having obtained the horizontal and vertical speeds of the eye, the speed (magnitude) and direction of the eye’s movement were calculated. This was done independently for each eye.

Saccades were detected independently in each eye using a velocity criterion of greater than 15 deg/s and a duration criterion of longer than 15 ms (cf. Erkelens & Sloot, 1995). An additional constraint on saccade detection was that the direction of eye movement at each point in the saccade was within ± 45 deg from the previous point.

Fixations were defined as periods of the eye movement record that were not saccades. Monocular portions of binocular saccades were defined as saccadic eye movements in which one eye moved with saccadic speeds prior to movement onset in the other eye. The eye movement records for each subject for each trial were checked to confirm that the computed analysis of each trial matched the actual eye position traces. In some cases, “monocular saccades” were identified in the eye movement recordings. These were observed when the movement of one eye was classified as saccadic but was below the saccadic threshold in the other eye. Monocular saccades were included in the analysis with binocular saccades after inspecting the eye position traces manually to confirm sub-saccadic movement in the other eye. In no instance did we observe genuinely monocular saccades where one eye moved rapidly and the other remained still (monocular viewing case described by Enright, 1992).

Observators

Three naive observers aged between 28 and 33 years took part in the experiment. Observers were naive to the purposes of the experiment, and two had not previously acted as observers in eye movement experiments. Observer KP has a slight, uncorrected astigmatism; CR is myopic but did not require eye correction for the viewing distances involved in the experiment; and, MB has normal vision.

Results

Eye Movement Strategies Under Different Task Constraints

First we examined the strategic nature of observers’ eye movements when performing the two different tasks by considering the locations and durations of observers’ fixations.

We started by presenting data on the gross differences between observers’ performance. These are evident from considering the average number and average duration of fixations presented in Table 1. It is clear that subjects made more fixations when performing the DEPTH task than the RATIO task.

Table 1. Average Number and Duration of Fixations When Performing the Two Tasks.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Task</th>
<th>Fixation count</th>
<th>Fixation duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td>MB</td>
<td>DEPTH</td>
<td>4.13</td>
<td>2.50</td>
</tr>
<tr>
<td>MB</td>
<td>RATIO</td>
<td>12.42</td>
<td>5.95</td>
</tr>
<tr>
<td>KP</td>
<td>DEPTH</td>
<td>4.77</td>
<td>3.72</td>
</tr>
<tr>
<td>KP</td>
<td>RATIO</td>
<td>8.13</td>
<td>2.83</td>
</tr>
<tr>
<td>CR</td>
<td>DEPTH</td>
<td>6.00</td>
<td>1.41</td>
</tr>
<tr>
<td>CR</td>
<td>RATIO</td>
<td>10.00</td>
<td>3.28</td>
</tr>
</tbody>
</table>

Fixation durations were defined as periods between gaze changes or gaze adjustments.

A smaller number of fixations in the DEPTH task might simply indicate that the task was easier, so fewer eye movements were made. However, consideration of fixation durations suggests that eye movement behavior was different in the two tasks. For observers KP and MB, fixations are considerably longer for the DEPTH task than the RATIO task (matched-pairs t tests on each subjects’ data; MB: t10=4.45, mean difference = 0.39s, p<.001; KP: t10=2.96, mean difference = 0.45s, p<.012). This difference was apparent from the raw eye movement traces: these subjects spent most of a trial keeping their eyes fixed on one or two of the patches in the scene. This was in marked contrast to their behavior when they performed the RATIO task where they made frequent saccades back and forth between the blocks. Subject CR showed the opposite pattern from the other two subjects. His fixations were slightly shorter in the DEPTH task than in the RATIO task. A Wilcoxon signed-ranks test (used due paired t test assumption violations) suggested a significant difference between the RATIO and DEPTH tasks (Z=2.31, p<.021).

For a more fine scale consideration of eye movement behavior, we now consider where observers fixated when performing the two tasks. Figure 3 shows the amount of time spent by observers looking at the three different patches when they were performing the two tasks.

From Figure 3 it can be seen that all observers, on average across trials, spend most time during a trial fixating the central patch. The effect of fixation duration was highly significant (general linear model on each observers’ data with patch location and task as factors; MB: F2,75=8.18, p<.001; KP: F2,84=6.79, p<.001; and CR: F2,60=6.55, p<.003). Post hoc Bonferroni comparisons revealed differences only between the fixation duration on the center patch and the left and right patches for all subjects. A significant interaction between the effects of task and patch location was observed for observers MB.
Welchman & Harris

and KP (MB: $F_{2,75}=3.82$, $p=.026$; KP: $F_{2,84}=3.78$, $p=.002$; and CR $F_{2,60}=0.80$, $p=.46$), indicating that the relative amount of time spent fixating the central patch differs between the two tasks. This is strikingly apparent for observers MB and KP; when they performed the DEPTH task, they spent the vast majority of their time fixating the central patch.

The general linear model also suggested that the effect of task was significant for two observers (MB: $F_{1,75}=5.61$, $p=.02$; KP: $F_{1,84}=2.57$, $p=.11$; and CR $F_{1,60}=37.37$, $p<.001$). To allow a comparison between tasks despite duration differences, we present the duration of gaze at the left and right patches relative to the duration of gaze at the central patch in Table 2.

Table 2. Relative Mean Time Spent Fixating Each Patch.

<table>
<thead>
<tr>
<th>Subject</th>
<th>RATIO task Left</th>
<th>RATIO task Right</th>
<th>DEPTH task Left</th>
<th>DEPTH task Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>MB</td>
<td>0.87</td>
<td>0.77</td>
<td>0.22</td>
<td>0.15</td>
</tr>
<tr>
<td>KP</td>
<td>0.79</td>
<td>0.70</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>CR</td>
<td>0.58</td>
<td>0.73</td>
<td>0.49</td>
<td>0.51</td>
</tr>
</tbody>
</table>

We used the mean duration of gaze on the central patch as a baseline to calculate the relative proportion (ratio) of fixation durations at the other patches (e.g., mean left duration divided by mean center duration).

The data from Table 2 highlight the point that observers performing the DEPTH task fixate the central patch proportionately longer than they do when performing the RATIO task. This effect is most striking for observers MB and KP. Indeed, the strategy of concentrating on the central patch was apparent from their raw eye traces. Whilst performing the DEPTH task for 5 trials out of 13 for observer KP, and for 5 trials out of 15 for observer MB, the observers spent the whole trial with their gaze directed to the central patch. This behavior was not seen in either subject when they performed the RATIO task. The strategy of fixating the central patch, and making judgments about the position of the other patches with respect to it, accords to the idea of fixating the notional reference point in the scene (cf. Foley, 1980).

Saccadic Changes of Vergence

Here we examine changes of vergence across saccades, and in particular whether there is a special role for isovergence saccades. For isovergent saccades to be useful, they would have to be made between the three different patches presented in the scene. For this reason, we divided our analysis of the vergence changes mediated by saccades into two groups: changes of gaze (defined as saccades having an amplitude of $> 3$ deg, corresponding to saccades between the different patches) and adjustments of gaze (defined as saccades of an amplitude $< 3$ deg, movements within a patch).

In Figure 4, we show a histogram of the changes of vergence mediated by saccades for gaze adjustments and changes of gaze for the two tasks. Gaze adjustments are associated with small changes of vergence (left-hand side of the figure), and the distributions for the two tasks are very similar (blue and red columns follow a similar pattern). A Chi-square test of homogeneity suggested no differences between the two distributions for each subject (MB: $\chi^2(3)=1.72$, $p>.25$; KP: $\chi^2(3)=4.31$, $p<0.25$; and CR: $\chi^2(3)=1.86$, $p>.25$). For gaze changes (right-hand
Figure 4. Vergence changes effected by saccades in the two tasks. The histograms show the relative frequencies of changes in vergence across saccades in the two tasks for all three observers. The left-hand column shows changes of vergence associated with gaze adjustments (saccade amplitude < 3 deg), and the right-hand column shows vergence changes made by changes of gaze (saccade amplitude > 3 deg). Blue bars show data from the RATIO task and red bars data from the DEPTH task.

Table 3. Number of Saccades Classified as Isovergent in the Two Tasks.

<table>
<thead>
<tr>
<th>Subject</th>
<th>RATIO task</th>
<th>DEPTH task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Isovergent</td>
<td>Total</td>
</tr>
<tr>
<td>MB</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>KP</td>
<td>2</td>
<td>64</td>
</tr>
<tr>
<td>CR</td>
<td>2</td>
<td>74</td>
</tr>
</tbody>
</table>

Data show the number of instances in which the change of vergence associated with a saccade was < 0.2 deg for all saccades made between stimulus patches that necessitated a change of vergence. We chose < 0.2 deg as our isovergence criterion because this was the vergence resolution of the eye tracker that we employed. Data presented in this table differ from those in Figure 4 because those data also included saccades made between stimulus patches requiring the same vergence angle.
RATIO task (KP: $\chi^2(1)=11.76, p < .001$; and CR: $\chi^2(1)=13.53, p < .001$), as would be expected if subjects exploited isovergent saccades to obtain useful information about depth structure. However, no differences were evident for observer MB ($\chi^2(1)=0$), who made a considerable proportion of isovergent saccades in both tasks. Whilst the proportion of isovergent saccades is larger in the DEPTH task for two of the observers, these saccades account for only a quarter of all the saccades made. This does not suggest that observers consistently and regularly locked the vergence state of their eyes across saccades.

Are There Task Specific Effects on Gaze Shifts?

Here we examine whether the observers’ task influences saccadic properties. To do this we consider the main-sequence relationship, the relationship between amplitude and duration, and the distribution of saccade amplitudes.

In Figure 5, the main sequence parameters for the three observers involved in the study under the two task conditions are shown. Saccade velocity is plotted as a function of saccade amplitude. Saccades made in the two tasks have very similar dynamics: a linear regression on the data for the two tasks in each subject suggested very similar regression coefficients (see Figure 5 legend). We find no evidence that main sequence parameters are different when subjects perform different tasks. These results contrast with those of Epelboim et al. (1997), who reported that higher peak velocities were observed for the same amplitude of movement according to the observer’s task.

A related saccadic property examined by Epelboim et al. (1997) was the duration of gaze shifts. They reported shorter saccadic durations when observers performed the TAP task. We examined gaze-shift duration as a function of gaze-shift amplitude and, in line with Figure 5, found no effect of the observers’ task (data not shown). Unlike Epelboim et al. (1997), for both main sequence parameters and gaze-shift durations, we found no systematic task effects.

In contrast, one difference between the eye movements produced by observers in the two tasks that can be appreciated from Figure 5 is that the distribution of saccade amplitudes produced by observers is not the same. Note that there are generally more red square symbols (DEPTH task) to the left of the figure. This is more clearly appreciated in Figure 6 where we present a histogram of the relative frequencies of the binocular saccade amplitudes for each task. The shape of the distributions in the two tasks is clearly different. This difference between the distributions of saccade amplitudes for the two tasks was confirmed as significant using tests of homogeneity (Chi-square tests on frequency counts of saccade amplitude [bin size 1.5 deg]: MB: $\chi^2(9)=64.59, p < .001$; KP: $\chi^2(8)=16.94, p < .05$; and CR: $\chi^2(8)=16.38, p < .05$).
When observers performed the RATIO task, they made a majority of saccades with an amplitude of 5 to 7 degrees, corresponding to the amplitude required to alternate gaze position between the central patch and the side patches. In contrast, when performing the DEPTH task, they produced a greater proportion of saccades with small amplitude (1 to 3 deg) than when they performed the RATIO task (i.e., the distribution shows a larger positive skew). Such eye movements correspond to adjustments of gaze as the observer re-fixates individual targets. This pattern of fixation behavior again appears consistent with the idea of fixating the notional reference point in the scene (cf. Foley, 1980).

**Binocular Saccade Dynamics**

Finally, we report on the nature of the integration of changes of vergence with changes of version. A number of studies over the past 20 years have suggested that vergence changes are often achieved by saccadic eye movements (e.g., Ono et al., 1978; Enright, 1984, 1986, 1992; Collewijn et al., 1988a, 1988b), allowing vergence changes to occur much more rapidly than by smooth disjunctive “pure vergence” eye movements (van Leeuwen, Collewijn, & Erkelens, 1998). Here we report similar findings for unpracticed, naïve observers making self-paced eye movements whilst engaged in a task unrelated to eye movement.

In Figure 7, we show representative mean binocular saccade traces from subject KP whilst she performed the RATIO task. The subject moved her eyes between the center patch and the right-hand patch (Figure 7A) and from the right patch to the left patch (Figure 7B).

When a change of vergence was required (Figure 7A), the observer produced saccades of different magnitudes in each eye: producing rapidly changing vergence angle across the saccade. When the observer made a saccade between targets requiring the same vergence angle (Figure 7B), there was transitory divergence of the eyes during the saccade followed by increased convergence at the end of the saccade. These observations of binocular saccadic properties concur with those reported elsewhere under more constrained task conditions (Enright, 1984; Collewijn et al., 1998a; van Leeuwen et al., 1998).
Discussion

In this work, we examined the nature of binocular eye movements made by observers when performing two rather different visual tasks in order to examine the role and function of binocular eye movements when making decisions based on visual information. Our analysis consisted of first examining the location and duration of fixations under different task constraints, and then examining whether the task performed by a subject influences saccade dynamics. Finally, we considered the nature of the binocular coordination of the eyes in naïve, unpracticed observers. We now review and discuss our findings.

Observers’ Strategic Eye Movements

In the “Introduction,” we suggested that three distinct patterns of eye movement could be produced by observers when making judgments requiring knowledge of the depth arrangements in the scene. First, observers might make regular isovergence saccades; second, observers might choose to fixate a reference point in the scene relative to which other points in depth are judged; and third, observers might simply make saccades in which changes of version and vergence are combined.

To evaluate the suggested eye movement patterns, we considered the location and duration of fixations. When performing the DEPTH task, observers spent predominately more time fixating the central patch (Figure 3) and, for two of the observers, fixation durations were considerably longer (Table 1). It was notable that the observers had different strategies for performing the DEPTH task. Observers MB and KP performed the DEPTH task by devoting long fixations to the central patch. In some cases, they did not move their eyes to the other patches throughout the whole trial. The strategy of fixating the central point and then keeping the eyes still appears compatible with Foley’s (1980) theory that observers choose a notional reference point in the scene relative to which disparities are scaled. Two of the observers may have chosen to fixate this notional reference point. By keeping the eyes still, it is possible to combine an estimate of distance to the fixated point with relative binocular disparity of other points with respect to
it, to give information about their relative distance apart in depth.

When performing the RATIO task, we suggested that control of vergence need only be sufficient to prevent diplopia (as information about depth structure was not required to perform the task). We observed that all the subjects performed the task in the optimal way (focusing each patch in turn to obtain the most accurate data about each one’s aspect ratio), and the eyes were generally well coordinated, associating changes of vergence with changes of version (Table 3).

We suggested that observers performing the DEPTH task might make isovergent saccades to register foveal absolute disparities rather than peripheral relative disparities, thus facilitating their judgments of spatial layout. Such behavior was observed by Enright (1991) when he ensured that relative disparities of two objects were not simultaneously available. We considered the magnitude of the vergence changes effected by the saccades under the two different tasks when observers moved their eyes between targets requiring a change of vergence. The results differed between observers (Table 3). Subject MB was found to make a number of isovergent saccades when saccades requiring a change of vergence were required (approximately 25%). This was true for both of the tasks we examined. Observers CR and KP showed task-dependent differences in the production of isovergent saccades: few saccades could be classified as isovergent in the RATIO task (ca. 3%); however, around a quarter of saccades were classified as isovergent in the DEPTH task. These results might suggest that observers adopted a strategy of isovergence to maximize disparity measurements. However, it should be noted that observers did not make a series of isovergent saccades in quick succession, as might be expected for such an “isovergence” strategy. On the basis of these data, it would be difficult to conclude that there was a special functional role for isovergent saccades when observers made depth judgments. It appears therefore that although observers can make isovergent saccades, they are probably not doing so specifically in order to measure absolute retinal disparity. If they are, then only a quarter of the saccades made are relevant to the task.

**Task Effects on Eye Movements**

We considered the mechanistic properties of eye movements to determine whether the high-level cognitive task being performed by the observer influences gaze-shift dynamics (cf. Epelboim et al., 1997). The data presented (Figure 5) suggested very little difference between the two tasks our observers performed. One potential explanation for this is that the saccades our observers produced had smaller amplitudes than some of those measured in the Epelboim study. The task-dependent effects in her study were most noticeable for large amplitude saccades, so perhaps we did not have saccades of sufficiently large amplitude to see an effect. A more likely explanation for our failure to find task-dependent differences comes from the fact that our observers were not free to move when performing the tasks. There were two sources of body movement in the Epelboim study. First, observers coordinated the movements of their eyes with the movement of their heads (Herst et al., 2001).

Accompanying movements of the head can lead to faster gaze-shift dynamics (Collewijn et al., 1992), possibly through changes in the gain of the vestibular ocular reflex (VOR). Observers performing the TAP task in the Epelboim study kept their heads less still and made faster head movement (Herst et al., 2001); if VOR gain was lower for faster movements (i.e., compensation for the movement of the head in the same direction was less accurate), then faster movements of the eyes might be expected compared with the LOOK task where the head moved less.

Another possible source for the increase in saccade speeds and a reduction in saccade durations could come from the fact that observers in the Epelboim study made accompanying movements of their arms in the TAP task. Snyder, Calton, Dickinson, and Lawrence (2002) recently examined the influence of arm movement on saccade parameters in rhesus monkeys. They found higher peak velocities and shorter saccadic durations when monkeys accompanied eye movements with arm movements to nearby targets. As the monkeys’ heads were stabilized, changes in VOR gain could not be responsible. They suggested that interactions between neural circuits planning eye and arm movement in the parietal cortex and/or interactions between neurons controlling movement parameters in the superior colliculus could lead to enhancement effects, whereby peak velocity is increased and duration decreased.

As the observers in this study moved neither their heads nor their hands, neither of the two tasks we employed would benefit from changes in VOR gain or “neural enhancement” effects. Thus task-related differences in saccade dynamics might arise only when the task the observer performs leads to other changes in the behavior of the subject (e.g., body movement), rather than being related to higher level cognitive demands placed on the use of visual information.

**The Binocular Coordination of the Eyes**

Our understanding of the way in which observers coordinate changes of vergence with changes of version is predominantly informed by experiments that have had observers instructed to repetitively look back and forth between small targets at regular intervals. Whilst the stimuli we employed cannot be described as natural, our observers were naïve, unpracticed, and were free to move their eyes as they liked. We observed binocular eye movements consistent with reports presented elsewhere (Enright, 1984; Collewijn et al., 1988a; van Leeuwen et
al., 1998): observers generally moved their eyes to combine changes of vergence with changes of version (Figure 7). This skill did not appear to require training. There were some exceptions however – observers did sometimes make isovergent saccades when changes of vergence angle were required (Table 3). We know of only one report of the extent of binocular coordination of the eyes when observers freely perform real world activities (Steinman, Epelboim, Forofonova, & Bogacz, 1998), which also appears to suggest that the oculo-motor system produces rapid binocular coordination of the eyes that allows us to move our eyes to the correct azimuth and the correct depth.

Summary

Everyday human activity routinely involves exploiting the same visual information for different purposes. Here we considered a simplified case of using the same visual information for different purposes. We observed that the use to which visual information is put can have a strong influence on the patterns of fixations, whilst not influencing saccade dynamics. By examining observers’ fixational strategies, we provide some support for the notion that observers choose and fixate a notional reference point in the scene, but find limited evidence for the use of isovergent saccades in making decisions about metric depth structure.

Acknowledgments

This research was supported by a grant from the Engineering and Physical Sciences Research Council (UK) to J.M.H. (GR/N13890). Commercial relationships: none.

References


