Available online at www.sciencedirect.com

SCIENCE DIRECT®

Progress in Retinal and Eye Research 25 (2006) 296–324

Progress in
RETINAL AND EYE RESEARCH

www.elsevier.com/locate/prer

Eye movements and the control of actions in everyday life

Michael F. Land

Department of Biology and Environmental Science, University of Sussex, Brighton BN1 9QG, UK

Abstract

The patterns of eye movement that accompany static activities such as reading have been studied since the early 1900s, but it is only since head-mounted eye trackers became available in the 1980s that it has been possible to study active tasks such as walking, driving, playing ball games and ordinary everyday activities like food preparation. This review examines the ways that vision contributes to the organization of such activities, and in particular how eye movements are used to locate the information needed by the motor system in the execution of each act. Major conclusions are that the eyes are proactive, typically seeking out the information required in the second before each act commences, although occasional 'look ahead' fixations are made to establish the locations of objects for use further into the future. Gaze often moves on before the last act is complete, indicating the presence of an information buffer. Each task has a characteristic but flexible pattern of eye movements that accompanies it, and this pattern is similar between individuals. The eyes rarely visit objects that are irrelevant to the action, and the conspicuity of objects (in terms of low-level image statistics) is much less important than their role in the task. Gaze control may involve movements of eyes, head and trunk, and these are coordinated in a way that allows for both flexibility of movement and stability of gaze. During the learning of a new activity, the eyes first provide feedback on the motor performance, but as this is perfected they provide feed-forward direction, seeking out the next object to be acted upon.

Contents

E-mail address: M.F.Land@sussex.ac.uk.

1350-9462/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.preteyeres.2006.01.002

1. Introduction

1.1. The need for eye movements during action

Throughout the animal kingdom, in animals with as diverse evolutionary backgrounds as men, fish, crabs, flies and cuttlefish, one finds a consistent pattern of eye movements which can be referred to as a 'saccade and fixate' strategy (Land, 1999). Saccades are the fast movements that redirect the eye to a new part of the surroundings, and fixations are the intervals between saccades in which gaze is held almost stationary. As Dodge showed in 1900, it is during fixations that information is taken in: during saccades we are effectively blind.

In humans there are two reasons for this strategy. First, the fovea, the region of most acute vision, is astonishingly small. Depending on exactly how it is defined, its angular diameter is between 0.3° and 2°, and the foveal depression (fovea means pit) covers only about 1/4000th of the retinal surface (Steinman, 2003). Away from the foveal centre resolution falls rapidly (Fig. 1). To see detail in what we are looking at, we need to move the fovea to centre the target of interest. Because a combination of blur and active suppression causes us to be blind during these relocations we have to move the eyes as fast as possible, and saccades are indeed very fast, reaching speeds of $700^\circ\,\mathrm{s}^{-1}$ for large saccades (Carpenter, 1988). Second, gaze must be kept still between saccades, during the fixations when we take in visual information. The reason for this is that the process of photoreception is slow: it takes about 20 ms for a cone to respond fully to a step change in the light reaching it (Friedburg et al., 2004). The practical effect of this is that at image speeds of greater than about 2–$3^\circ\,\mathrm{s}^{-1}$ we are no longer able to use the finest (highest spatial frequency) information in the image (Westheimer and McKee, 1975; Carpenter, 1991): in short, the image starts to blur, just as in a camera with a slow shutter speed. Interestingly, animals without well-defined foveas still employ the saccade and fixate strategy. Keeping gaze rotationally stable is the primary requirement whatever the retinal configuration, but in addition mobile animals necessarily require saccadic gaze-shifting mechanisms. Without such a mechanism, when the animal makes a turn the eyes will

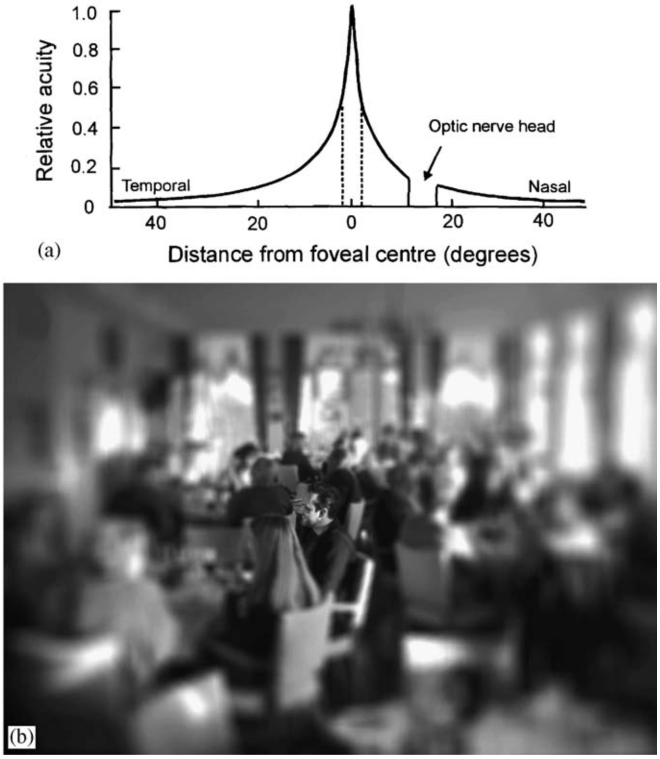


Fig. 1. (a) Relative grating acuity across the horizontal visual field. This has already fallen to half its maximal value by the edge of the fovea (2° across: dotted lines). Based on data from Wertheim (1894). (b) Consequences for the image transmitted by the retina of the loss of resolution with eccentricity. Radially increasing blur, corresponding to the acuity decrease in (a), has been added to a photograph. The picture represents approximately the central 20° of the field of view. Picture by Ben Vincent. From: Basic Vision, an Introduction to Visual Perception. Oxford University Press (2006).

counter-rotate until they become stuck at one end of their movement range (Walls, 1962).

During ordinary activity, the body and head rotate the eyes in space at velocities as high as several hundred degrees per second, so that for fixation to be maintained during such motion powerful compensatory mechanisms are required to move the eyes in the opposite direction to the rotation of the head. These mechanisms are of two kinds. In the vestibulo-ocular reflex (VOR) the semi-circular canals measure head rotation velocity, and the signal they provide is fed to the eye muscles via the vestibular and oculomotor nuclei. The gain of this reflex is close to 1, so that a rotation of the head evokes an eye movement that almost exactly counteracts it. At slower velocities a second reflex, the optokinetic reflex (OKR), takes over from VOR. It operates by measuring the actual velocity of the image on the retina, and causes the eye muscles to rotate the eye in the same direction as the retinal motion, thus nulling it out. OKR is a feedback system, working on the error between the desired image speed ($0\,^{\circ}\,s^{-1}$) and its actual speed. VOR on the other hand is not a feedback mechanism, as the movements of the eyes have no effect on the sensor—the semi-circular canals. Between them these two reflexes keep eye rotation in space within acceptable limits. Residual image motion, under conditions of realistic natural head rotation is in the range 0.5–$5\,^{\circ}\,s^{-1}$ (Collewijn et al., 1981; Kowler, 1991), i.e. close to the limit at which blur would start to set in.

In ordinary active life, these two types of eye movement—saccades and stabilizing movements—dominate. Two others are important and need to be mentioned. Small moving objects can be tracked by the smooth pursuit system. Here the target is kept on the fovea by smooth movements not unlike those of OKR. However, OKR operates on large areas of the image whereas pursuit requires a small target, and when a target is being tracked the pursuit system is actually pitted against the wide-field OKR system, whose function is to keep the overall image still. Smooth pursuit on its own only works up to target velocities of about $15\,^{\circ}\,s^{-1}$. Above this speed the smooth movements are supplemented by saccades, and above about $100\,^{\circ}\,s^{-1}$ pursuit is entirely saccadic. Vergence movements are responsible for adjusting the angle between the eyes to different distances, and they are unique in that the eyes move in opposite directions relative to the head. The role of vergence in real tasks is unclear. In principle, the eyes should converge so that the two foveal directions intersect at the target, but during a task where the subjects had to use vision to guide tapping, 'vergence tends to be set 25–45% *beyond* the attended plane; in other words, subjects do *not* adjust gaze to intersect the attended target' (Steinman, 2003, p. 1350). It may well be that, out of the laboratory situation, vergence control is quite imprecise.

These, then, are the components from which eye movement strategies in real life tasks are constructed. They are essentially the same as those studied under various kinds of restraint in laboratories over the past century. There are other issues that have been less well studied in laboratory conditions, for example the co-operative actions of eye, head and body, which become important in the behaviour of freely moving individuals. And we may not necessarily expect the same constraints on eye movements outside the laboratory as we find when subjects are asked to do their best at some artificial task. To quote Steinman (2003, p. 1350) again: 'Under natural conditions gaze control is lax, perhaps even lazy. One could just as easily call it *efficient*. Why should the oculomotor system set its parameters so as to make it do more work than is needed to get the job done'.

1.2. Recording eye movements and fixation strategies: a brief historical overview

Objective studies of human eye movements date from around the turn of the twentieth century, although methods involving the use of after-images go back to the 18th century (Wade and Tatler, 2005). The first eye movement recordings were made by Delabarre in 1898, using a mechanical lever attached to the eye via a plaster of Paris ring (!). Dodge and Cline (1901) introduced a method for photographing movements of the reflection of a light source from the cornea, which remained the standard method of recording eye movements for 50 years (Steinman, 2003). The method was used in various forms, notably by Buswell (1920) to study reading aloud, and later to record eye movements made while looking at pictures (Buswell, 1935). Butsch (1932) used it to study eye movements during copy typing, and the eye movements of pianists during sight-reading were examined by Weaver (1943). The method required the head to be kept as still as possible, because any head movement changes gaze direction relative to the object being viewed, and so makes it impossible to determine where the eye is looking from eye-in-head movements alone. Improvements of the technique by Ratliff and Riggs (1950) permitted a modest amount of head movement (by using a collimated beam to put the object being observed at infinity), but nonetheless eye movement recordings were still limited to subjects who were essentially stationary. This meant that the study of the kinds of eye movements made during most of the active tasks of everyday life was precluded.

The first devices that made it possible to record eye movements during relatively unconstrained activity were made by Mackworth and Thomas (1962). They used a camera mounted on the head—they had both cine and TV versions—which simultaneously filmed the view ahead and the corneal reflection. By means of some ingenious optics they combined the images so that the moving dot produced by the corneal reflection was superimposed on the scene view to give the location of foveal gaze direction (Thomas, 1968). In this way they could visualize directly where the eye was looking, and because the device was head-mounted the 'problem' of head movement no longer existed. The device was used successfully to study both driving and flying. However, it was heavy and not particularly accurate (about 2° visual angle), and the design was not taken up by others. For a time another recording method seemed promising: the use of search coils mounted on both the eye and head (search coils generate currents when they rotate in the magnetic field of a larger surrounding coil). The combined output gives gaze direction, or eye and head direction separately (Collewijn, 1977). However, wearing a search coil for any length of time is uncomfortable, and movements are only possible within the magnetic field of the external coils, and the method has only been used in laboratory situations. By the 1980s video cameras had become much smaller and lighter, and a number of commercial eye trackers, along the lines of the Mackworth and Thomas cameras, began to become available. They were usually based on pupil position, made visible by illuminating the eye with infra-red light to produce a 'white' pupil which is tracked electronically. Its location relative to the head is then transferred as a spot or crosshair to the image from the scene camera, to give a pictorial display of foveal gaze direction, or 'point of regard' (Duchowski, 2003). These eye trackers are now in common use (Fig. 2). A variant tracks the iris rather than the pupil (Land, 1993), and many of the records in this review were made with such an arrangement. Head-mounted eye trackers, in combination with an external video camera to record motor activity, are the main tools required to explore the relations between eye movements and motor actions.

It is appropriate here to mention two studies that have had a profound effect on the development of the field. The first was by the Russian physiologist Alfred Yarbus. He recorded the eye movements of subjects looking at pictures, extending the earlier work of Buswell (1935). Yarbus got his subjects to look at the pictures with a number of different questions in mind (Yarbus, 1967). These might relate to the relationships of the people in the picture, or the clothes they were wearing (see Section 2.1.4 and Fig. 5). What he found was that each question evoked a different pattern of eye movements, clearly related to the information required by the question. This meant that eye movements were not simply related to the structure of the picture itself, but also to 'top-down' instructions from executive regions of the brain. The significance of this for the present review is that when we are engaged in some activity, such as carpentry or cookery, we are also presented with a series of questions—'Where is the hammer?' 'Is the kettle boiling?'—which can only be answered if appropriate eye movements are made. Yarbus' work provided the precedent for abandoning the older idea that eye movements were basically reflex actions, and demonstrated that they are much more strategic in character.

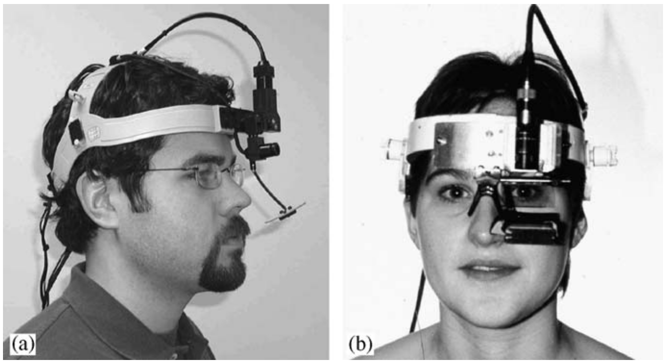


Fig. 2. Recent head-mounted eye-tracking cameras. (a) Device in which the eye is illuminated with infra-red light to produce a bright pupil which is tracked on the video image. A second camera captures the scene ahead. Courtesy of Applied Science Laboratories (a-s-l.com), (b) Device in which the outline of the iris is tracked on the video image. This uses only one camera with a split field of view. This prototype was used for many of the recordings discussed in this review (Land, 1993).

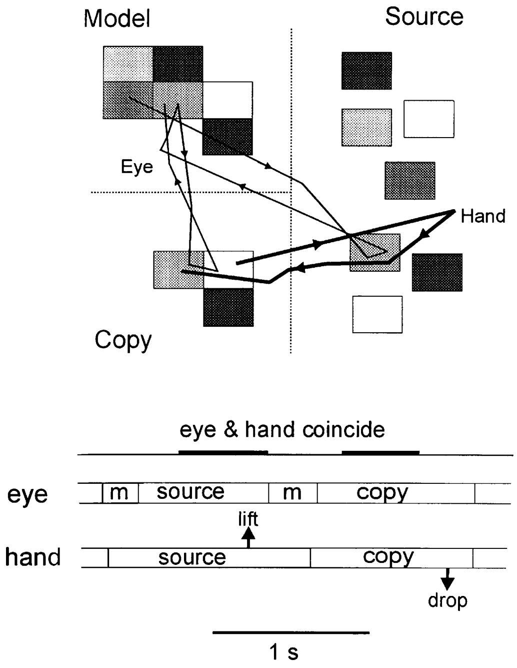


Fig. 3. Block copying task devised by Ballard et al. (1992). Upper part shows the layout of the task, with the eye movements (thin line) and hand movements (thick line) made during one cycle. Lower section shows the timing of eye and hand locations during an average cycle. See text for details. Modified from Ballard et al. (1992).

The second study is really the one that ushered in the present era of exploring the relationship between eye movements and actions. Ballard et al. (1992) devised a task in which a model consisting of coloured blocks had to be copied using blocks from a separate pool. Thus the task involved a repeated sequence of looking at the model, selecting a block, moving it to the copy and setting it down in the right place (Fig. 3). The most important finding was that the operation proceeds in a series of elementary acts involving eye and hand, with minimal use of memory. Thus a typical repeat unit would be as follows. Fixate (block in model area); remember (its colour); fixate (a block in source area of the same colour); pickup (fixated block); fixate (same block in model area); remember (its relative location); fixate (corresponding location in model area); move block; drop block. The eyes have two quite different functions in this sequence: to direct the hand in lifting and dropping the block, and, alternating with this, to gather the information required for copying (the avoidance of memory use is shown by the fact that separate glances are used to determine the colour and location of the model block). The only times that gaze and hand coincide are during the periods of about half a second before picking up and setting down the block.

The main conclusion from this study was that the eyes look directly at the objects they are engaged with, which in a task of this complexity means that a great many eye movements are required. Given the relatively small angular size of the task arena, why do the eyes need to move so much? Could they not direct activity from a single central location? Ballard et al. (1992) found that subjects could complete the task successfully when holding their gaze on a central fixation spot, but it took three times as long as when normal eye movements were permitted. For whatever reasons, this strategy of '*do it where I'm looking*' is crucial for the fast and economical execution of the task. As we shall see, this strategy seems to apply universally. With respect to the relative timing of fixations and actions, Ballard et al. (1995) came up with a second maxim: the '*just in time*' strategy. In other words, the fixation that provides the information for a particular action immediately precedes that action; in many cases the act itself may occur, or certainly be initiated, within the lifetime of a single fixation. It seems that memory is used frugally here, as testified by the fact that separate fixations are used to obtain the colour and relative position of the blocks (although in other tasks memory for object location can persist for quite long periods, as we shall see later in Section 3.1.1). The conclusions from these studies are substantially borne out by most of the examples detailed in part 2 of this review, and they can be regarded as basic rules for the interaction of the eye movement and action systems.

1.3. *Scope of this review: questions to be addressed*

This review differs from most previous reviews of eye movements (e.g. Carpenter, 1988) in that it is not concerned with eye movements per se, but rather with the functions of the sequences of fixations that accompany different kinds of activity. The latter part of the twentieth century saw a huge amount of experimental work devoted to the physiology of eye movements. This included the mechanics and neuromuscular physiology of the eye, the nature of the control systems involved, and the neurophysiology of the central mechanisms responsible for their generation (see Robinson, 1968, 1981; Carpenter, 1988, 1991). Much recent effort has gone into working out how different regions of the brain—in particular, the superior colliculus—are involved in the generation of saccades (Gandhi and Sparks, 2003; Sommer and Wurtz, 2003). At the same time much psychological research has gone into saccade generation, especially in the fields of attention and visual search (Findlay and Gilchrist, 2003; Schall, 2003). Almost all these studies deal with eye movements as single entities: saccades, stabilizing reflexes, pursuit and vergence were mainly considered as isolated systems rather than components of a larger strategy (although work on search patterns comes closest to this). It is this larger strategy—how we use our eyes to obtain the information that we need for action—that I will address here, and I will not deal in any great detail with the individual components, whose characteristics are well reviewed elsewhere.

Different kinds of activity have different requirements for visual information. A tennis player has to assess the trajectory of a rapidly approaching ball in order to formulate a return stroke. A pianist needs to acquire notes continuously from the two staves of a score, translate them into finger movements and emit them simultaneously as a continuum of key strokes. A driver must simultaneously keep the car in lane, avoid other traffic and be aware of road signs. A cook following a recipe must perform a succession of acts of preparation and assembly, each one different from the others, in a defined sequence. In all of these activities, the eyes provide crucial information at the right time and from the right place, and the patterns of fixations are unique to the particular task.

The rest of the review is in two main parts. In Section 2, I will present descriptions of the patterns of eye movements and fixations that accompany different types of activity. This will provide a data base which I will mine in Part 3 to address some of the questions that the different studies throw up. For example: What kinds of information do the eyes supply to the motor system of the limbs? How close does gaze have to be to the site of the action it is controlling? When is visual information acquired and supplied in relation to the timing of the motor actions themselves? What does the oculomotor system need to know about the location of objects in order to find the appropriate information? How do eyes, head, limbs and trunk cooperate in the production of an action? What can we learn about the central mechanisms responsible for these patterns of coordination? What role does memory play? Except in the context of reading, and some other sedentary activities, few of these questions were addressed prior to about 1990, and many of them remain unanswered.

2. Examples of fixation strategies and their relations to action

2.1. *Sedentary activities*

The eye movements associated with activities in which the head could be kept still were amenable to study from the time of the very earliest eye movement recordings. For example Erdmann and Dodge (1898; see also Dodge, 1900) first showed that during reading the subjectively smooth passage of the eye across the page is in reality a series of saccades and fixations, in which information is taken in during the fixations.

2.1.1. Reading

Although silent reading involves no overt action, it nevertheless requires a particular eye movement strategy to make possible the uptake of information in a way that allows meaning to be acquired. It is also one of best studied (as well as the most atypical) examples of a clearly defined eye movement pattern. Eye movements in reading are highly constrained to a linear progression of fixations to the right (in English) across the page, which allows the words to be read in an interpretable order. In this reading differs from many other activities (such as viewing pictures) where order is much less important (Buswell, 1935). Reading is a learned skill, but the eye movements that go with it are not taught. Nevertheless, they are remarkably similar between normal readers. Eye movements during reading have been reviewed thoroughly recently, and only the principal facts need to be included here. Most of what follows is derived from an extensive review by Rayner (1998). The reader is referred to Radach et al. (2004) for accounts of recent issues in reading research.

During normal reading, gaze (foveal direction) moves across the line of print in a series of saccades, whose size is typically 7–9 letters. Within limits this number is not affected by the print size, implying that the oculomotor system is able to make scaling adjustments to its performance. For normal print the saccade size is 1–2°, and the durations of the fixations between saccades have a mean of 225 ms. When reading aloud fixations are longer (mean 275 ms). Most saccades (in English) are to the right, but 10–15% are regressions (right to left) and are associated in a poorly understood way with problems in processing the currently or previously fixated word. Words can be identified up to 7–8 letter spaces to the right of the fixation point, but some information is available up to 14–15 letter spaces; this is used in the positioning of subsequent saccade end points. From studies in which words were masked during fixations, it appears that the visual information needed for reading is taken in during the first 50–70 ms of each fixation. Adult readers typically read at about 300 words per minute or 0.2 s per word.

If changes are made to the text during the course of a fixation, both the duration of the current fixation, and the size of the following saccade can be affected. This implies that the text is processed 'on-line' on a fixation by fixation basis. Similarly, difficult words result in longer fixations, indicating that cognitive processes operate within single fixations. How long it takes to process words all the way from vision to meaning is hard to assess. However the delay between reading and speech during reading aloud (the eye-voice span) can be measured (Fig. 4a). In a classic study Buswell (1920) found that high-school students had an eye-voice span of about 13 letters, or 0.79 s (average word length of 4.7 letters and a reading speed of 3.5 words per second). On a simpler reading piece elementary school students (5th grade) had an eye-voice span of 11 letters, or 0.91 s. The eye-voice span of 'good' and 'poor' readers differed by 4–5 letters.

2.1.2. Music reading

Musical sight-reading shares with text reading the constraint that gaze must move progressively to the right. It is, however, more complicated in that—for keyboard players—there are two staves from which notes must be acquired. Weaver (1943) recorded eye movements of trained pianists, and found their gaze alternated fixation between the upper and lower staves (Fig. 4c). 'Notes on the treble and bass parts of the great staff are usually so far apart that both vertical and horizontal movements of the eyes must be used in preparing two parallel lines of material for a unified performance' (Weaver, 1943, p. 27). This alternation means that notes that have to be played together are viewed at different times, adding a task of temporal assembly to the other cognitive tasks of interpreting the pitch and length of the notes. For the Bach minuet illustrated in Fig. 4c, Weaver's pianists acquired notes from the score at between 1.3 and 2.0 notes per fixation (making a note roughly equivalent to a word in text reading). Interestingly, the fixations on the upper stave were much longer (average 0.44 s) than those on the lower stave (0.28 s), presumably because more notes were acquired during each upper stave fixation. The time from reading a note to playing it (the eye–hand span) was similar to what Buswell (1920) had found for reading aloud: for the minuet the average for 10 performances was 3.1 notes, or 0.78 s. Furneaux and Land (1999) looked at the eye–hand span in pianists of differing abilities. They found that it did not vary with skill level when measured as a time interval, but that when measured in terms of the number of notes contained in that interval professionals averaged four compared with two for novices. Thus the processing time is the same for everyone, but the throughput rate of the processor is skill dependent. The processing time did vary with tempo, however, with fast pieces having an eye–hand span of 0.7 s, increasing to 1.3 s for slow pieces.

2.1.3. Typing

Copy typing, like music playing, has a motor output, and according to Butsch (1932) typists of all skill levels attempt to keep the eyes about 1 s ahead of the currently typed letter, which is much the same as in music reading. This represents about five characters (Fig. 4b). More recently Inhoff and his colleagues (Inhoff and Wang, 1992) found more variability in the eye–hand span, and also showed that it was affected by the nature of the text. Using a moving widow technique they showed that typing starts to become slower when there are fewer than three letter spaces to the right of fixation, indicating a perceptual span about half the size of that used in normal reading. The potential word buffer is much bigger than this, however. Fleischer (1986) found that when typists use a read/check cycle of

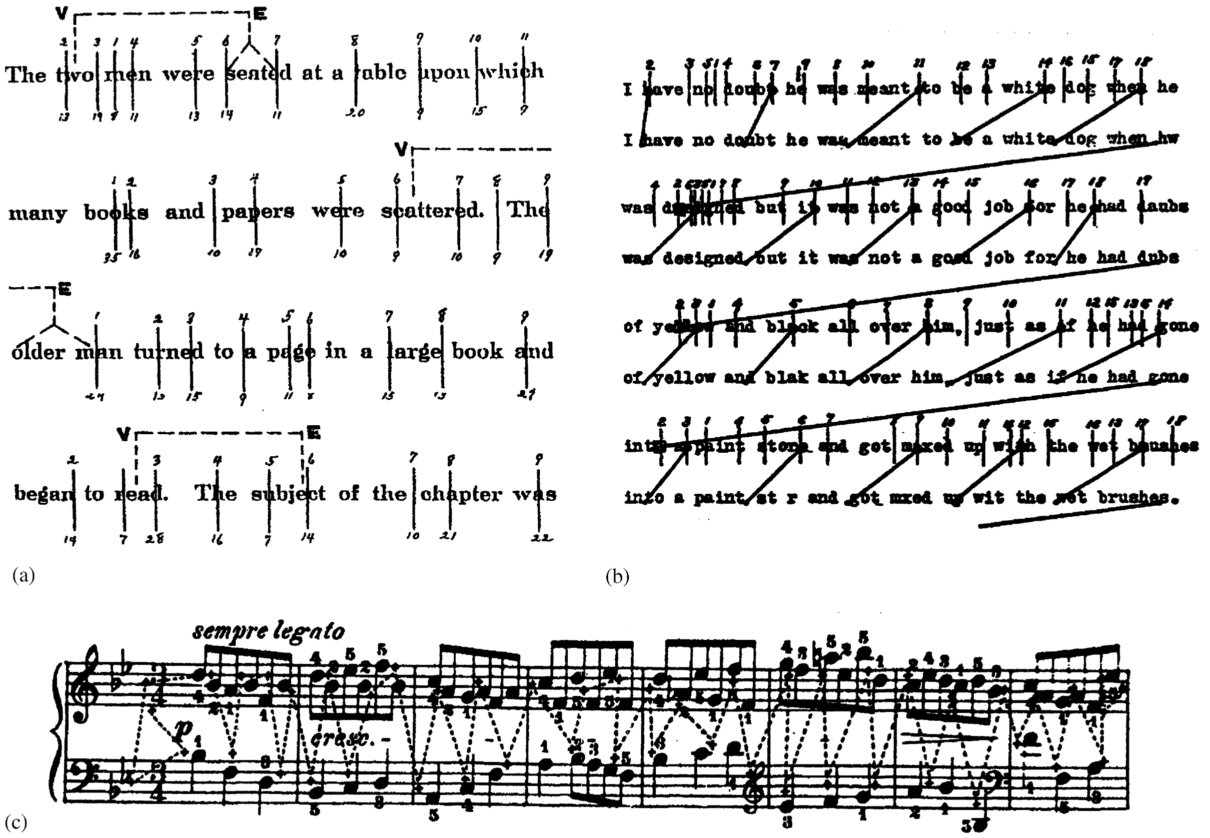


Fig. 4. Classic recordings of eye movements during various sedentary tasks. (a) Eye-voice span (V–E) during reading aloud by a high-school student (freshman), recorded in 1920 by Guy Buswell. Vertical bars are fixations: upper numbers give the sequence along each line, and lower numbers their durations in multiples of 0.02 s. V–E indicates the eye position at the time sound was uttered. (b) Record of typing by Butsch (1932) showing the fixation sequence (upper lines: numbers give fixation order) and the line actually typed (lower lines) including errors. Oblique lines show eye positions at the instants the keys were pressed. (c) Record by Weaver (1943) of the eye movements of a pianist playing a Bach minuet. Dotted lines join successive fixations, and show how gaze alternates between the staves.

approximately 1 s each, whilst typing continuously, they would typically take in 11 characters during the read part of the cycle, and exceptionally strings of up to 30 characters could be stored.

These three activities—reading, musical sight-reading and typing—are all similar in that they involve the continuous processing of a stream of visual information taken in as a series of stationary fixations. This information is translated and converted to a stream of muscular activity of various kinds (or into 'meaning' in the case of silent reading). In each case the time within the processor is about a second. Once the appropriate action has been performed, the original visual information is overwritten so that the operation is more like a production line than a conventional memory system.

2.1.4. Looking at pictures

Viewing a picture has no obvious output, and the patterns of eye movement are rather less constrained than in the preceding activities. Nevertheless Buswell, in his 1935 book 'How people look at pictures' found that patterns of eye fixations were related to the structures in the pictures, albeit in a rather loose way. Two of his conclusions were particularly interesting. First, he observed that fixation patterns changed during the viewing period, with later fixations being of longer duration than earlier ones, and spread out more across the picture. He suggested that the viewer changed from an initial quick survey to a more detailed study of limited regions. This finding was repeated by Antes (1974) who found an increase in fixation duration from 215 ms initially to 310 ms after a few seconds. Recent studies of scene viewing also propose a distinction between the first fixations on the scene, which tend to be similar between viewers, and later fixations where viewers' strategies diverge, driven more by the meaning, or semantic interest, of regions of the scene (Henderson and Hollingsworth, 1998; Tatler, Baddeley and Gilchrist, 2005). Second, Buswell found that asking subjects particular questions about the picture changed the patterns of the eye movements. He showed subjects a photograph of the Tribune tower in Chicago. In one trial the eye movement record 'was obtained in the normal manner without any special directions being given. After that record was secured, the subject was told to look at the picture again to see if he could find a person looking out of one of the windows of the tower.' (Buswell, 1935, p. 136). The pattern of eye

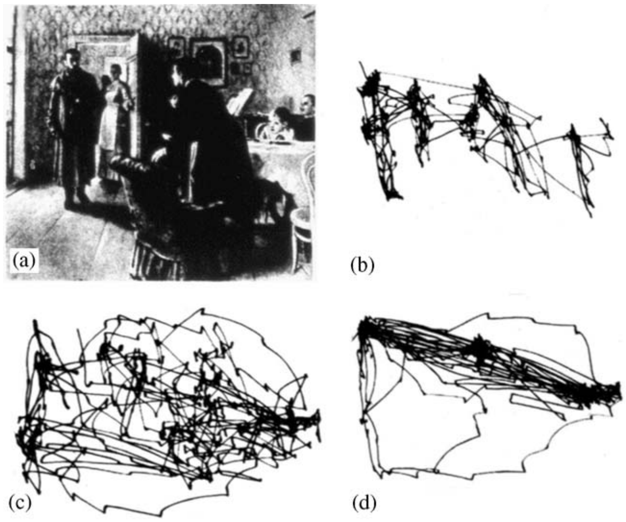


Fig. 5. Recordings made by Alfred Yarbus of the eye movements a subject viewing a picture ('They did not expect him' by I.P. Repin) with different questions in mind. (a) The picture. (b) 'Remember the clothes worn by the people'. (c) 'Remember the positions of the people and objects in the room'. (d) 'Estimate how long the 'unexpected visitor' had been away.' Modified from Yarbus (1967).

movements in the second case was quite different, with a much greater concentration of fixations around the regions with windows. This was actually a very important observation, because it demonstrated for the first time that eye movements are not just triggered in a reflex way by the conspicuity of the objects in the scene, but are also subject to 'top-down' control by instructions related to the demands of particular tasks.

The studies of Alfred Yarbus (1967) have already been mentioned in Section 1.2. He demonstrated this kind of top-down control even more impressively. He made recordings of the eye movements of subjects as they viewed a number of pictures, and with one particular picture, 'They did not expect him' by I. Repin (Fig. 5a), he asked the subjects a series of different questions. Yarbus used a method in which a mirror was attached to the subject's eye with a suction cup, and a beam of light reflected from the mirror wrote directly onto a photographic plate. Fig. 5 shows three of his records. In (b) he asked the subjects to remember the people's clothes, which produced vertical saccades, in (c) to remember the positions of the people and objects, a task so large as to produce something like oculomotor panic, and in (d) to estimate how long the 'unexpected visitor' had been away. This involves finding subtle clues to emotions from faces, and almost all fixations are indeed on the faces. Not only is the role of top-down instructions apparent in the pattern of fixations, but so also is the near absence of fixations on objects that are not relevant; for example in (c) the maid who opened the door is not fixated, as her face cannot help to answer the question. In much of what follows Yarbus' insights are the starting point, as they lead on to the question: 'What happens if it is not the investigator who asks the questions, but the participant himself?' The performance of any activity requires that the visual system interrogate the surroundings with a series of questions about the presence, locations and states of the objects that the task entails.

2.1.5. Drawing and sketching

The task of producing a picture is very different from than simply looking at one. In drawing a portrait the artist has to acquire information from a sitter, formulate a line to be drawn and execute this on the drawing itself. There is thus a repeating sitter—drawing gaze cycle, with vision employed in different ways in each half cycle. In the first study of its kind Miall and Tchalenko (2001) recorded the eye and hand movements of the portrait artist Henry Ocean as he made first a pencil sketch of a model (12 min) and then a finished drawing (100 h over 5 days). In both cases there was a very regular alternation between sitter and drawing, with 0.6–1.1 s spent on the sitter and rather longer (2–4 s) on the drawing. Ocean's fixations on the sitter were always single, whereas novice artists who were also studied spent shorter periods on the model but often made multiple fixations. Miall and Tchalenko estimate that while drawing a line Ocean was capturing about 1.5 cm of detail per fixation on the model, and that visual memory was being refreshed roughly every 2 s (Tchalenko et al., 2003).

A problem in trying to probe deeper into the way vision is used in the production of each line during a long drawing session is that the functions of each sitter-drawing cycle are not all the same. Sometimes a line is drawn, sometimes it is just checked, sometimes a line is altered or added to, and often, particularly in Henry Ocean's portrait, a line is simply rehearsed without the pencil contacting the paper. One way to remove this ambiguity is to make a fast sketch in which there is no checking, and a line is drawn every cycle. In our laboratory we asked a painter and art teacher, Nick Bodimeade, to make some portrait sketches for us, as well as a longer more measured drawing, whilst wearing an eye tracker with a head-mounted scene camera that showed both the sitter and the drawings (M.F. Land and G. Baker, hitherto unpublished study). Fig. 6a shows the whole sequence for one sketch, together with the 'average cycle' in which the various timings are indexed to the beginning of each drawn line. The principal findings were that a typical cycle lasted 1.7 s (35 cycles per minute), with 0.8 s on the sitter and 0.9 s on the sketch (Fig. 6b). On average the pen made contact with the paper about 0.1 s after gaze transferred to the sketch, and lasted for the time gaze remained on the sketch. However there was much variation, as Fig. 6a shows, and standard deviations for all these measures (relative to the beginning of the drawn line) were in the range 0.3–0.5 s, so the cycles were far from metronomic, and no event was absolutely synchronized to any other.

It was possible to work out something of what was happening as the artist formulated his next line. Between one and four fixations were made on the sitter's face

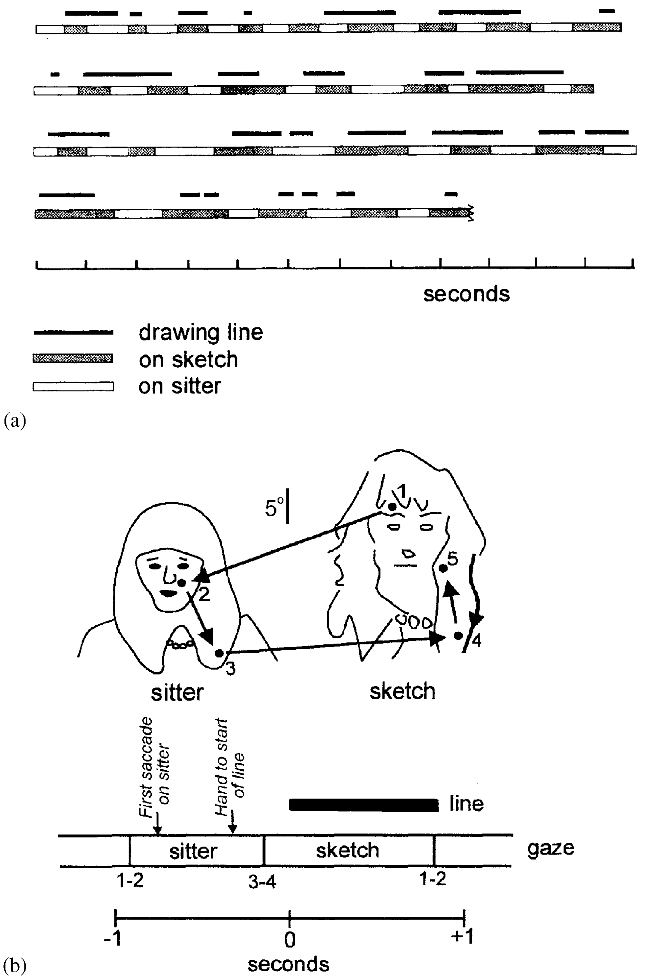


Fig. 6. Eye–hand strategy of an artist making a 40 s portrait sketch. (a) Alternation of eye movements between sitter and sketch, and their time relations with the lines drawn. (b) Events during an average drawing cycle, derived from data in (a). Explanation in text.

(mean 2.3), and by the last fixation the point to be addressed on the sketch had been selected. When gaze left the sitter, it was transferred accurately (<2° error) to the corresponding point on the sketch. Interestingly, this was not the point that the next line was to be drawn *from*, but the point drawn *to,* i.e. the end of the line (Fig. 7). This surprised both ourselves and the artist. It does, however, make some sense. In a sketch each line is a new entity, almost unrelated to the last. Thus start of the next line must be determined by some process of internal selection by the artist. (This contrasts with the detailed drawings made by both Nick Bodimeade and Henry Ocean, where one line usually continued on from or was closely related spatially to its predecessor.). The course of the line and its end-point, however, are derived from features on the sitter, once the start of the line has been established.

The selection of the target point (i.e. the first fixation on the sketch and the endpoint of the next line) occurred during the first fixation on the sitter, which was unusually short (~0.15 s). Subsequent fixations were longer (~0.28 s), but did not bring gaze closer to the target. Interestingly, when only one fixation was made on the sitter it was of long duration (~0.43 s), equal to the sum of the first and second fixations when two were made. We speculate that in this case it takes 0.15 s to make the decision that the gaze is already on target, and that the function of the rest of that fixation, and of subsequent fixations when more than one is made, is to obtain information about the form of the line to be drawn. The timing of the selection of the position of the start of the line is more problematic, because the first sign of hand movement to the start point does not occur until about half a second after the end point is established (Fig. 6b). However, it seems logical that the beginning and end of the line would be determined at about the same time. In contrast to the more measured drawing, the last

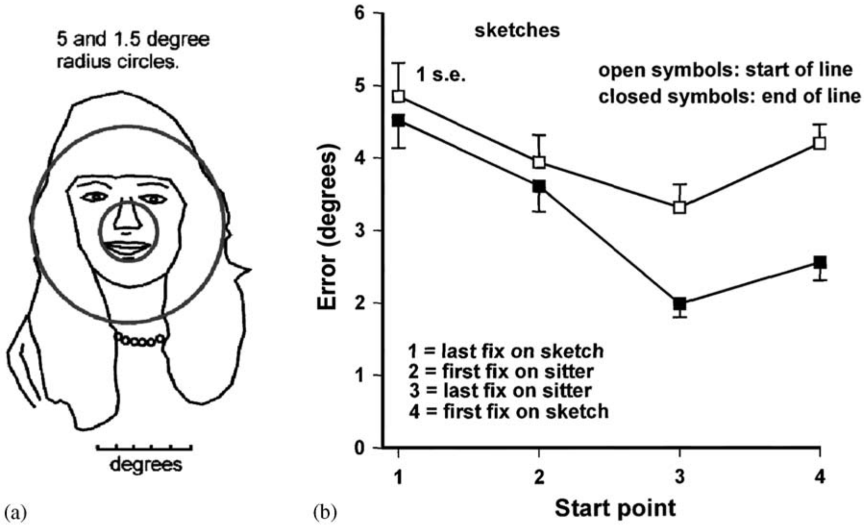


Fig. 7. Locations of fixations on the sitter and sketch in relation to the line about to be drawn. Numbers on abscissa refer to the numbered positions on Fig. 6b. Error on the ordinate is the angular distance (on the sketch) between the fixations and either the beginning or end of the next line. The change in error while on the sitter (2 and 3) shows that the artist selects the end of the next line, not its start point. Insert on left shows the meaning of the error scale in relation to the sitter's face.

fixation on the sketch was a very poor predictor of either the start or end point of the next line, so it seems that all decisions about the position and shape of the line to be drawn are made while gaze is on the sitter. There was no evidence from the sketches of strategic planning beyond the next line.

It is worth noting that this kind of sketching is essentially a copying task, and can be compared with the block copying task of Ballard et al. (1992). The timings of the repetitive cycles, and the components within them, are remarkably similar (Figs. 3 and 6b).

2.2. Locomotion: walking and stepping

The study of the gaze movements of head-free, fully mobile subjects required the development of eye-trackers that were head—rather than bench-mounted. These were not available except as difficult-to-use prototypes until the 1980s (see Section 1.2), and so most of the results described below are products of the last 20 years. Most of the new generation of eye trackers provide a head-based view of the scene ahead, with the direction of regard (of the fovea) represented by a spot or crosshair. Frequently, the motor behaviour of the participants is also filmed.

When crossing level ground, walkers rarely need to look at where they are going to step safely. However, in more difficult terrain they tend to fixate the location of their future footfalls. An obvious question is how far they look ahead, in order to obtain the information they need for a safe footfall. This was addressed by Patla and Vickers (2003) who required their subjects to step on a series of irregularly spaced 'footprints' over a 10 m walkway. They found that subjects fixated the footprints on average two steps ahead, or 0.8–1 s in time. We have repeated these findings in Sussex (M. Armstrong, C. Isbell and M. Land, hitherto unpublished) using a damaged pavement on which the subjects were instructed not to tread on the cracks (Fig. 8). We used an eye tracker with a second synchronized camera directed at the feet. For five subjects the average number of steps between footfall and the nearest fixation to it was 1.91 (s.d. 0.53), and the average time lag 1.11 s (s.d. 0.46 s), very much in line with Patla and Vickers' result. As can be seen from Fig. 8 there are roughly two fixations per step, but there is no simple correspondence between fixation points and footfalls. Typically, the nearest fixation to a footfall is about 5° from it.

Using a light spot to indicate the location of 'undesirable' foot placements, Patla et al. (1999) found that the foot could be redirected to a new location within the duration of the preceding step, and that these alternative placements were not selected at random. They were generally directed to the location that minimized the displacement of the foot from what would have been its normal footfall, so causing the least disruption to both locomotor muscle activity and dynamic stability. It thus appears that footfalls are typically planned up to two steps into the future, but adjustments can be made within one step if required.

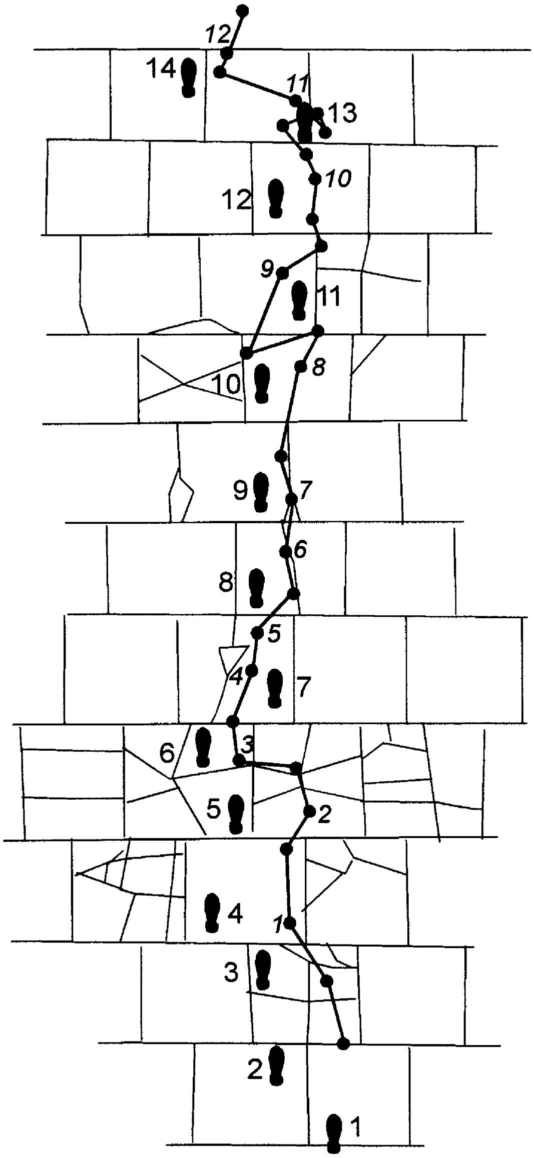


Fig. 8. Location of gaze (line and dots) and footfalls while walking across cracked paving, with the instruction not to step on the cracks. The numbers on the gaze track indicate the fixations that occur at the same time as the correspondingly numbered footfalls. On this record, gaze is typically two steps ahead of the footfall.

What happens when we change direction? Ultimately it is the body axis that rotates, carried by the feet, but what rôles do eye and head movements play? Hollands et al. (2002) studied this by getting subjects to change direction along a travel path whilst wearing eye and head monitoring equipment (this was flat terrain so there was no need to fixate future footfalls). The direction change was indicated either by the onset of a cue light at the new path end point, or by prior instruction about the route. The principal finding was that the turn was invariably accompanied by an eye saccade to the new destination (with a latency of about 350 ms when cued), initiated as the same time as a head movement. The eye–head combination brought head and gaze into line with the direction of the new goal as the body turn was being made. Thus gaze movements into the turn anticipate body movements, and the authors argue that pre-aligning the head axis provides an allocentric (external) reference frame that can then be used for the control of the rest of the body. Something very similar occurs when body turns are made without forward motion (see Land, 2004, and Section 3.4), and when turning a corner in a car (see Section 2.3.5).

2.3. *Driving*

Driving is a complex skill that involves dealing with the road itself (steering, speed control), other road users (vehicles, cyclists, moving and stationary pedestrians) and attention to road signs and other relevant sources of information. It is thus a very varied task, and one would expect a range of eye movement strategies to be employed. I will first consider steering, as this is a prerequisite for all other aspects of driving.

2.3.1. *Steering on winding roads*

When steering a car on a winding road, vision has to supply the driver's arms and hands with the information needed to turn the steering wheel by the right amount at the right time. What is this control signal, and how is it obtained?

Early studies, mainly on US roads that had predominantly low curvatures, had found only a weak relationship between gaze direction and steering (e.g. Zwahlen, 1993). In 1994, David Lee and I decided to look at the more visually demanding task of steering on a road whose bends were continuous and unpredictable. Queens Drive round Arthur's seat in Edinburgh was ideal—very winding, but one-way and so without the distraction of other traffic. We found a much clearer relationship between direction of gaze and steering. In particular, drivers spent much of their time looking at the 'tangent point' on the up-coming bend (Land and Lee, 1994; Underwood et al., 1999). The tangent point is the moving point on the inside of each bend where the driver's line of sight is tangential to the road edge; it is also the point that protrudes most into the road, and is thus highly visible (Figs. 9 and 10a). It moves around the bend with the car but—for a bend of constant curvature—remains in the same angular position relative to the driver's heading. The angular location of this point relative to the vehicle's line of travel (effectively the driver's trunk axis if he is belted in) predicts the curvature of the bend: larger angles indicate steeper curvatures. Thus, potentially, this angle can provide the signal needed to control steering. Fig. 10c does indeed show that records of gaze direction and steering wheel angle are very similar. The implication is that this angle, which is equal to the eye-in-head plus the head-in-body angle when the driver is looking at the tangent point, is translated more or less directly into the motor control signal for the arms.

The geometry of the tangent point in relation to the curvature of the bend is shown in Fig. 10b. The curvature of the bend ($1/r$, the reciprocal of the radius) specifies the angle through which the steering wheel needs to be turned to match the bend—at least at reasonable speeds. It is

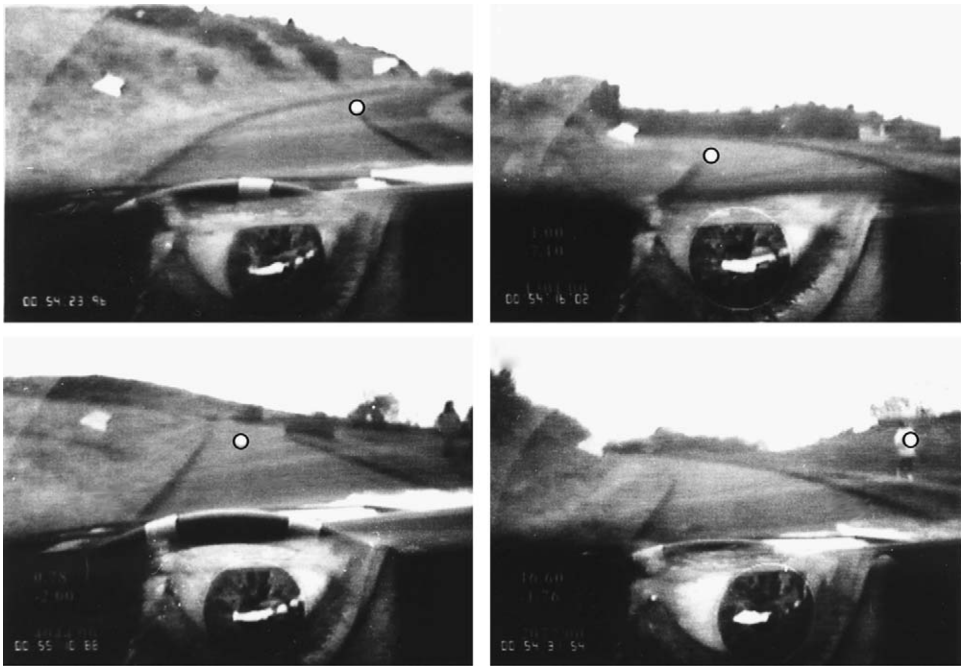


Fig. 9. Four views of a winding one-lane road showing typical gaze locations, taken from a video made with the device shown in Fig. 2b. Upper row shows right and left bends, with gaze directed to the tangent points; lower row shows typical gaze position on a straight road, and a glance off the road to look at a jogger. Upper part of each figure shows the view from the camera attached to the head; lower part shows the inverted eye imaged by a concave mirror. The position of the dot in the upper part is derived from the location of the outline of the iris. From Land (1998).

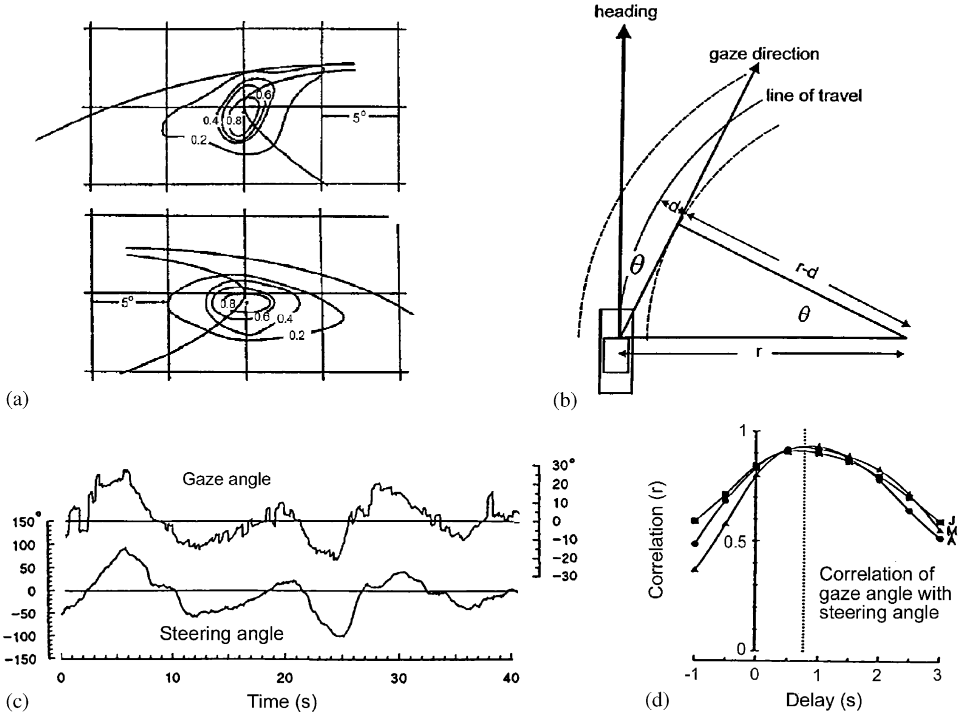


Fig. 10. (a) Contour plots of the relative density of fixations when driving round left and right bends on a narrow winding road. 60% of all fixations lay within the 0.2 contour. There is a strong peak within 1° of the tangent point in both cases. Three drivers, approximately 200 fixations per second (b) Geometry of the tangent point. The angle θ between the current heading and the tangent point provides a good measure of bend curvature ($1/r$). See text. (c) Relationship between gaze angle (θ on (b)) and the angle of the steering wheel, through a series of bends. Apart from the occasional fixation off the road, and a slight delay, the two records are nearly identical. (d) Cross-correlation between records like those in (b) for three drivers, showing the peak correlation occurs after a delay between gaze and steering angles of about 0.8 s. Modified from Land and Lee (1994) and Land (1998).

related to the gaze angle θ by

$$\cos\theta = (r - d)/r.$$

However, we can use the expansion: $\cos\theta \approx \theta^2/2$, which then gives

$$1/r = \theta^2/2d.$$

Thus the steering wheel angle is directly related to θ^2, and inversely related to d, the distance of the driver from the kerb, or inside lane edge on a multiple lane road. Evidently, in addition to measuring the angle θ, the driver must also either measure d, or else maintain it at a constant value. We will return to this point in Section 2.3.2.

It is important that the driver does not act on the tangent point signal immediately, because the part of the bend whose curvature he is measuring still lies some distance ahead. Cross-correlating the two curves in Fig. 10c shows that gaze direction precedes steering wheel angle by about 0.8 s (Fig. 10d), similar to eye-effector delay in other activities (see Section 3.1.3). This is much longer than a simple reaction time (typically around 0.3 s), and so represents an intended delay. This lag provides the driver with a reasonable 'comfort margin', but it is also the delay necessary to prevent steering taking place before a bend has been reached.

The tangent point is special in two other ways. It is a near stationary point in the velocity flow field: other points on both sides of the road move laterally in the visual field, and so will carry the driver's eye with them via the optokinetic reflex. The tangent point only moves when road curvature changes and this, as we have seen, is the signal the driver needs to steer by. Second, if the view around the bend is occluded, say, by a fence or hedge, then the tangent point affords the longest clear view of the road ahead. These various attributes make it unsurprising that tangent points are preferentially fixated (Fig. 10a). However, experience suggests that we are able to steer adequately without actually fixating the tangent point, for example when attending to road signs or other traffic. Fig. 10c and similar records show that the eyes are indeed not absolutely glued to the tangent point, but can take time out to look at other things. These excursions are accomplished by gaze saccades and typically last between 0.5 and 1 s. The probability of these off-road glances occurring varies with the stage of the bend that the vehicle has reached, and they are least likely to occur around the time of entry into a new bend. At this point drivers fixated the tangent point 80% of the time (Land and Lee, 1994). It seems that special attention is required at this time, presumably to get the initial estimate of the bend's curvature correct. A confirmation of this came from Yilmaz and Nakayama (1995), who used reaction times to a vocal probe to show that attention was diverted to the road just before simulated bends, and that sharper curves demanded more attention than shallower ones. The fewer and shallower the bends in the road, the more time can be spent looking off the road, and this probably accounts for the lack of a close relation between gaze direction and steering on studies of driving on freeways and other major roads. Nevertheless it is certainly true that looking away from the road ahead for any substantial period of time is detrimental. According to Summala (1998), lane keeping on a straight road deteriorates progressively when drivers are required to fixate at greater eccentricities from the vanishing point. There is only a slight drop in performance by 7°, this becomes substantial by 23°, and worse again by 38°. The effect is likely to be much more pronounced on bends, especially if the curvature changes. There are probably implications here for the positioning of both road signs and in-car controls.

Studies on a simple simulator have shown that feed-forward information from the distant part of the road is not on its own sufficient to give good steering (Land and Horwood, 1995). When the only the furthest region of the simulated road was visible, curvature matching was still accurate, but position-in-lane control was very poor (Fig. 11A). Conversely, with only the near-road region visible, lane maintenance was quite good, but curvature matching was poor, mainly due to rather wild 'bang-bang' steering induced by the short time (<0.5 s) available for reaction to the movements of the road edges (Fig. 11C). Although it would seem from Fig. 11B that somewhere in between, about 5° down from the horizon, gives a good result on both criteria, it turned out that the best performance was obtained when distant (A) and near (C) regions were combined. This was better than having region B on its own, and was indistinguishable from having the whole road visible. Interestingly the near part of the road was rarely fixated compared with the more distant region, but it was certainly seen and used; it is typically about 5° obliquely below the typical direction of gaze. Mourant and Rockwell (1970) had already concluded that lane position is monitored with peripheral vision. They also argue that learner drivers first use foveal vision for lane keeping, then increasingly move foveal gaze to more distant road regions, and learn to use their peripheral vision to stay in lane. Summala et al. (1996) reached similar conclusions. The principal outcome of these studies is that neither the far-road input (from tangent points), nor the near-road lane-edge input is sufficient on their own, but the combination of the two allows fluent accurate driving (Land, 1998).

2.3.2. *Models of steering behaviour*

As early as 1978 an engineer, Edmund Donges, showed that there are basically two sorts of signal available to drivers: feedback signals (lateral and angular deviation from the road centre line, differences between the road curvature and the vehicle's path curvature), and feed-forward or anticipatory signals obtained from more distant regions of the road up to 2 s ahead in time (corresponding to 90 ft or 27 m at 30 mph). Donges (1978) used a driving simulator to demonstrate that each of these signals was indeed used in steering, although he did not discuss how they might be obtained visually. A version of his steering model is shown in Fig. 12. It now seems that we can identify the feed-forward and feedback elements in the Donges scheme with the far-road (tangent point, vanishing point) and near-road (lane edge) visual inputs demonstrated in Fig. 11. It is not difficult to see why both are required. The far-road signal may provide excellent curvature information, but if the car starts out of lane, it will stay that way, however well it follows road curvature. One might think that lane-edge feedback on its own would be sufficient, but visual processing and mechanical delays mean that the feedback loop becomes unstable at even moderate speeds (as for example when driving in fog, with no far-road input). Matching road curvature takes the pressure off the near-road feedback loop, and it means that

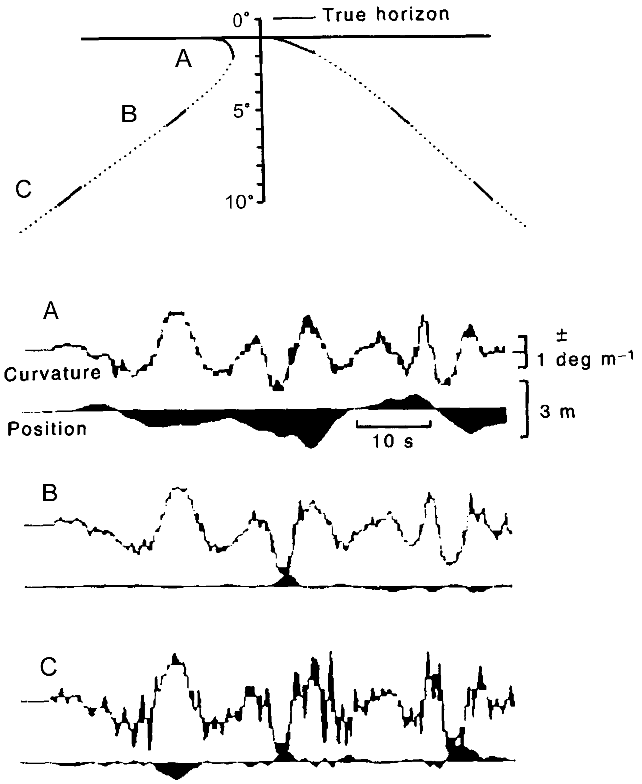


Fig. 11. Recordings of steering performance made using a rudimentary driving simulator in which most of the road edge was omitted except for 1° high segments at (A), (B) and (C). The bends in the road imitated those used for Fig. 10c. The upper part of each record shows the curvature of the vehicle's track in relation to the curvature of the road: a thickening of the line indicates a difference between the two. The lower part of each record is similar, but for position relative to the centre of the lane: a mismatch between car and road gives a thickened line. When only distant road regions are visible (A) curvature matching is good, but lane position maintenance is poor. With only near regions (C) lane position maintenance is acceptable, but curvature maintenance is unstable ('bang-bang') as the driver has difficulty coping with the feedback delay. Mid regions (B) give the best result on both measures, but other experiments show that a combination of distant and near regions (A and C) is even better. From Land and Horwood (1995).

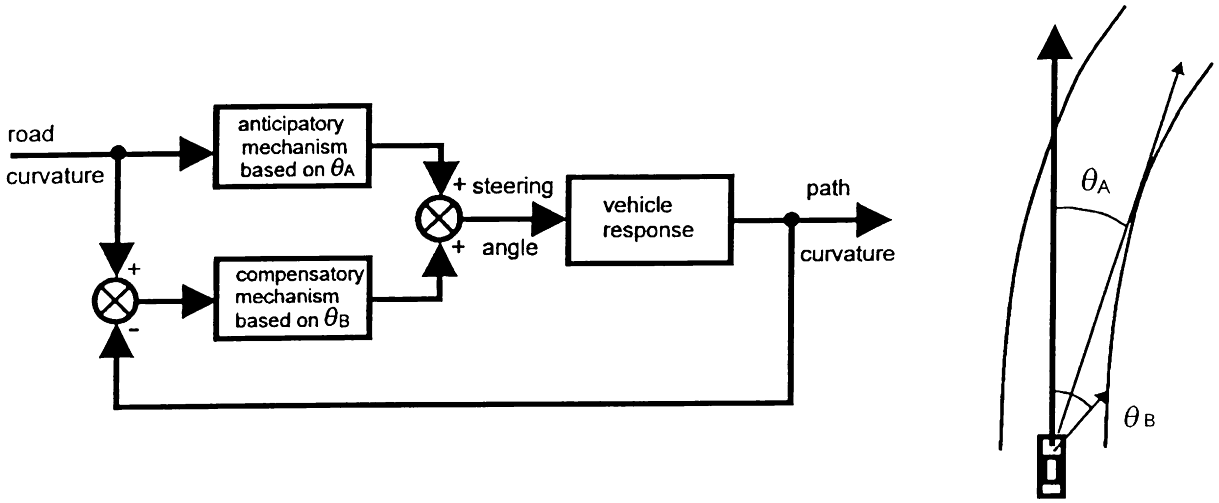


Fig. 12. Control diagram, based on Donges (1978), showing the combined use of anticipatory (feed-forward) information from distant road regions (θ_A) and feedback information from the near-road edge (θ_B).

it can operate at much lower gain, and is thus much less prone to instability.

2.3.3. Multitasking

Sometimes the eye must be used for two different functions at the same time, and as there is only one fovea and off-axis vision is poor, the visual system has to resort to time-sharing. A good example of this is shown in Fig. 13, in which the driver is negotiating a bend and so needs to look at the tangent point, while passing a cyclist who needs to be checked on repeatedly. The record shows that the driver alternates gaze between tangent point and cyclist several times, spending half a second on each. The lower record shows that he steers by the road edge, which means that the coupling between eye and hand has to be turned off when he views the cyclist (who would otherwise be run over!). Thus not only does gaze switch between tasks, so does the whole visual-motor control system. Presumably, whilst looking at the cyclist, the information from the tangent point is kept 'on hold' at its previous value in an appropriate buffer.

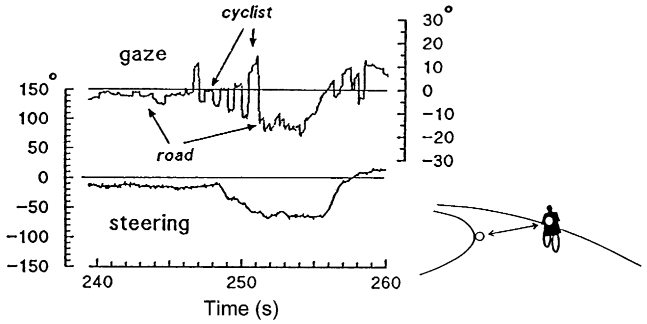


Fig. 13. Time sharing. Record of the gaze direction and steering wheel angle of a driver negotiating a bend while keeping an eye on a cyclist. Gaze alternates between tangent point and cyclist, with fixations of about 0.5 s on each. The fixations on the cyclist do not affect the steering, implying that activation of the steering and checking control systems also alternates with gaze.

2.3.4. Urban driving

Steering round the kind of right-angle corner we encounter in cities is a rather different task from following the curves of a country road. It is a well-rehearsed, rather stereotyped task, with the amount the steering wheel has to be turned varying little from one corner to another. The following account is based on a new study of three drivers each negotiating eight suburban right-angle corners. Each turn proceeds in two distinct phases, which, by analogy with ordinary walking turns, we can call orientational and compensatory phases (Imai et al., 2001). In the orientational phase gaze is directed into the bend by 50° or more relative to the car, with most of the rotation performed by the neck (head/car in Fig. 14); meanwhile the eyes fixate various positions around the bend. Once the car turn has begun the neck reverses its direction of rotation (seconds 53 and 9 in the two examples in Fig. 14), and the head starts to come into line with the car. However, it continues to rotate in space for a while, carried by the continued rotation of the car. This is the compensatory phase, so-called because the head rotation counteracts to a large degree the rotation of the car. As can be seen in Fig. 14 the car-in-space and head-in-car rotations are almost (but not quite) equal and opposite during this phase. This strongly suggests that the head is being stabilized by a feedback mechanism in which the vestibular system measures the residual head-in-space rotation, and converts it into a neck rotation command that counteracts the head-in-space rotation (Land, 2004). There is a known reflex, the vestibulo-collic reflex, which operates in just this manner. At about the same time as the neck reverses its direction of rotation, gaze shifts from the entrance to the bend to more distant regions of the road.

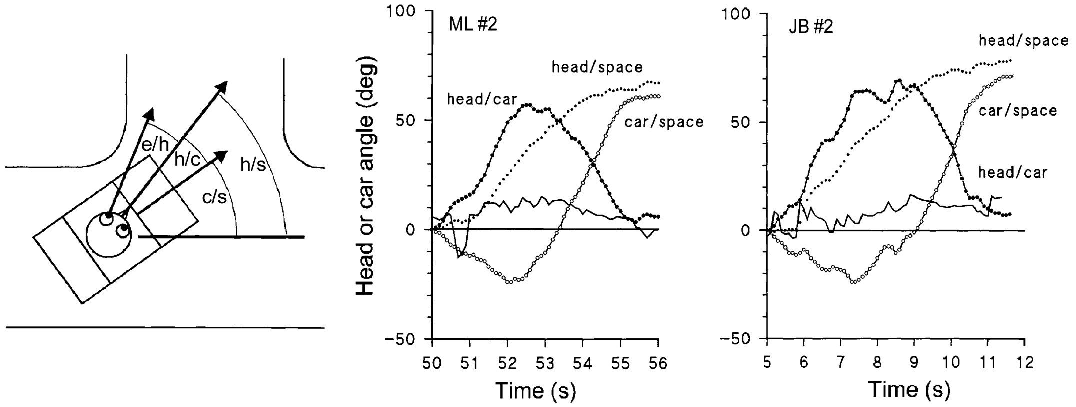


Fig. 14. Records of two drivers turning the same left-hand urban corner (near-side in UK). The principal feature of both records is the orientation of the head, which rotates into the bend during the first 3 s so that it leads the car's heading by as much as 70°. Thereafter as the car continues to turn, the neck rotates the head back into line with the car at a speed that is almost equal and opposite to the rate of rotation of the car itself. The effect of this compensatory rotation is that the head direction in space stays nearly constant, rotating by a further 20° as the car rotates through 70°. The effect of this manoeuvre is that gaze is directed to the exit of the bend almost as soon as turning has begun, and remains there throughout the turn. The driver is thus in a position to anticipate potential hazards several seconds ahead. Plain line shows the eye in head angle. See also Fig. 16.

What is critical in getting this manoeuvre right is the timing of the steering action, both when entering and exiting from the corner. Using the view provided by the eye-tracker, it was possible to examine what timing cues were available in the half-second or so before the driver began to steer into and out of the bend. The changes in the appearance of the road-edge (kerb) seemed to be the only cues to provide useful timing information, and which also correlated reliably with the initiation of the steering action (Fig. 15). In a left hand turn (nearside in the UK) the tangent point slips leftward as the corner approaches (angle α), and steering starts when α reaches between 30° and 40° (Fig. 16b and c left). The cue for straightening up at the exit of the bend seems to be rotation of the nearside kerb in the visual field (angle β). Just before the end of the bend the kerb angle rotates through the vertical in the driver's view, with β going from acute to obtuse (Fig. 16b and c right). The change of steering direction occurred when this angle reached between 140 and 150°, about half a second after the kerb passes through the vertical in the visual field. Although these may not be the only features involved, there was little else in the drivers' field of view that was both conspicuous and reliable. Turning right (offside in the UK) is a little more difficult as there are the added problems of crossing a traffic stream and lining up with the far kerb. However, similar cues are also available for this manoeuvre.

In urban driving, multi-tasking, is even more important than it is on country roads (cf. Fig. 13) as each traffic situation and road sign competes for attention and potential action. To my knowledge there has been no systematic study of where drivers look in traffic, but from our own observations it is clear that drivers foveate the places from which they need to obtain information: the car

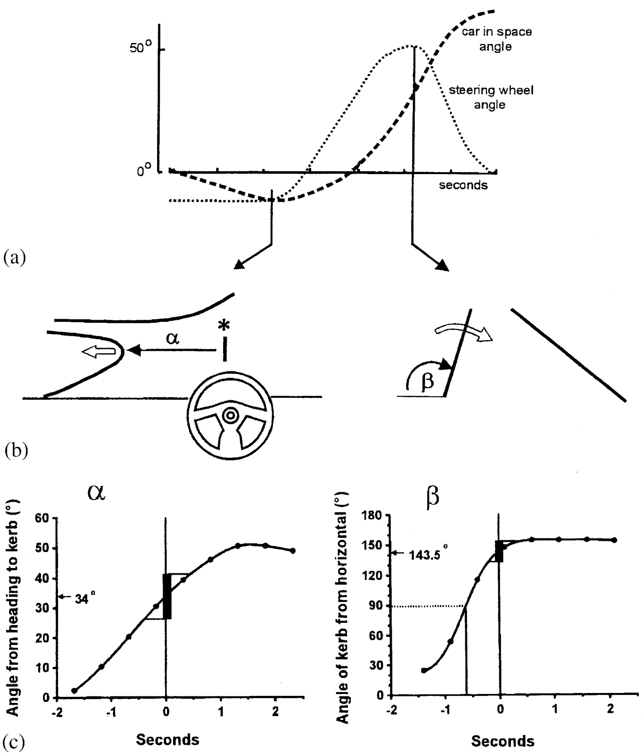


Fig. 15. Cues for the timing of steering action when negotiating a right-angle corner. (a) Average car rotation profile and steering wheel rotation for three drivers negotiating four left-hand (nearside) corners. (b) In the driver's visual field the most conspicuous cue for starting the turn into the corner is the increasing lateral position of the tangent point (angle α). At the exit from the turn the vertical rotation of the near-side kerb (angle β) provides a timing cue for reversing the steering wheel direction. (c) Values of α and β at the time that the steering wheel began to be turned left and right, respectively (see (a)). The vertical bars show the standard deviation of the initiation points for all 12 corners. The arrows on the ordinate show the mean values of α and β. Dotted lines on right-hand graph show the time at which the driver's view of the kerb passes through the vertical.

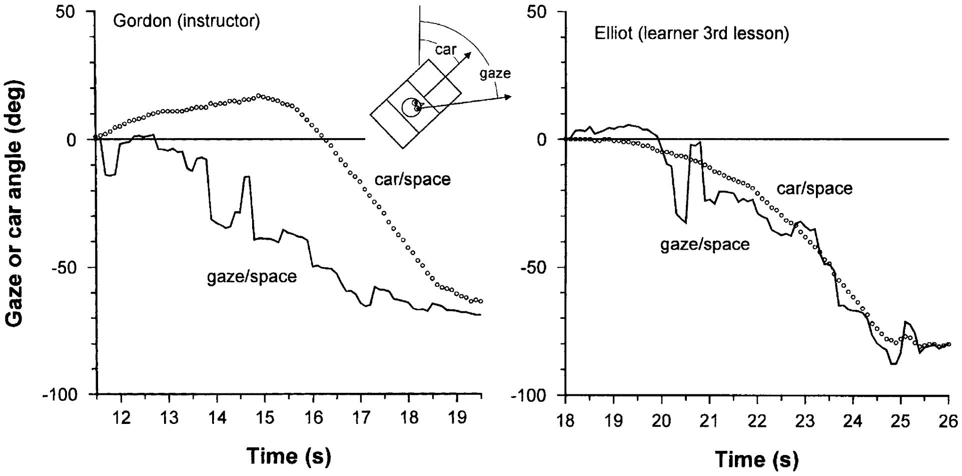


Fig. 16. Rotational trajectories of car heading (dots) and gaze (line) for a driving instructor (left) and a novice driver (right) during his third lesson. The instructor's gaze leads the car's heading by up to 50° at the start of the turn (or about 3 s in terms of the road ahead). The learner's gaze stays in line with the car throughout the turn.

in front, the outer edges of obstacles, pedestrians and cyclists, road signs and traffic lights, etc. In general, speeds of 30 mph or less appear only to require peripheral lane-edge (feedback) information for adequate steering. Thus the need to use distant tangent points is much reduced compared with open road steering, freeing up the eyes for the multiple demands of dealing with other road users and potential obstacles. Just as with open-road steering, both foveal and peripheral vision are involved. Miura (1987) has shown that as the demands of traffic situations increase, peripheral vision is sacrificed to provide greater attentional resources for information uptake by the fovea. Crundall (2005) also found that when potentially hazardous situations became visible in a video clip the ability of drivers to detect peripheral stimuli (lights) in the periphery was diminished. This effect was greater with novices compared with experienced drivers, and recovery (the time taken to re-engage peripheral detection) was also faster with the experienced group.

2.3.5. Learning to drive

In their first few driving lessons, learners have to master a great many unfamiliar patterns of sensory-motor coordination. These include steering (staying in lane, turning corners), braking, gear changing (exacerbated by the clutch in a stick-shift car), using the rear-view mirror, looking out for road signs, and staying vigilant for other vehicles and pedestrians. Initially all these tasks require attentive monitoring, but over the course of a few lessons many of them become at least partially automated, so that more attentional resources can be devoted to the less predictable aspects of driving, notably the behaviour of other road users.

There have been few studies of the very first stages of driving, if only because few learners are willing to add an unfamiliar eye tracker to their already onerous coordination tasks. However, in a recent study (M.F. Land and C.J. Hughes, hitherto unpublished) we did, with the consent of the police, use an eye tracker to examine the differences in gaze patterns between three novice drivers and their instructor, during their first four lessons. There were a number of minor differences relating to where the learners looked on straight roads, in particular they tended to confine gaze to the ahead direction with fewer glances off the road than experienced drivers. However the most striking and consistent effect was on the gaze behaviour of the novices when turning a corner (Fig. 16). The driving instructor, like most competent drivers, directed gaze by as much as 50° into the bend, soon after the steering wheel began to turn (see also Fig. 15). This was done almost entirely with a head movement; the eyes fixated various roadside objects towards the exit of the bend but rarely made excursions of more than 20° from the head axis. All three learners, on the other hand, kept gaze strictly in line with the car's heading, at least during the first lesson (by lesson four two of the three were turning into the bend like the instructor, but not the third). The significance of this change is that the novices are learning to anticipate: In Fig. 16, by second 15, the instructor is already looking at a point that the car's heading will not reach for another 2–3 s. Presumably this allows him to plan his exit from the bend and also notice whether there are any potential hazards. The learners cannot do this to begin with, probably because the task of getting the steering right for the bend requires all their attention. The reduced functional field of view, seen in novice drivers by a number of authors (Mourant and Rockwell, 1972; Crundall et al., 1998), is presumably also related to the fact that steering itself has yet to be fully mastered.

2.3.6. Racing driving

We have had one opportunity (Land and Tatler, 2001) to examine the eye and head movements of a racing driver (Tomas Scheckter) when driving at speed. Like ordinary drivers, his gaze was directed close to the tangent points of bends. However, unlike low-speed driving this was almost entirely the result of head rotation, rather than eye-in-head movements, which were of low amplitude ($<\pm 10^\circ$) and almost unrelated to the head movements. The most impressive finding, and one for which we have yet to find a convincing explanation, was that the angle of the head in the yaw plane was an almost exact predictor of the rotational speed of the car, 1 s later. Thus during a left hand hairpin, when the car was turning at $60^\circ\,\mathrm{s}^{-1}$ the head had turned 50° to the left 1 s earlier. It seems that the driver has in his brain a curvature map of the circuit which he uses to control his speed and racing line, but quite why this should manifest itself in the amount by which he turns his head is not at all clear.

2.4. *Ball sports*

Some ball sports are so fast that there is barely time for the player to use his normal ocular-motor machinery. Within less than half a second (in baseball or cricket) the batter has to judge the trajectory of the ball and formulate a properly aimed and timed stroke. The accuracy required is a few cm in space and a few ms in time (Regan, 1992). Half a second gives time for one or at the most two saccades, and the speeds involved preclude smooth pursuit for much of the ball's flight. How do practitioners of these sports use their eyes to get the information they need?

2.4.1. *Table tennis*

Part of the answer is anticipation. Ripoll et al. (1987) found that international table-tennis players anticipated the bounce and made a saccade to a point close to the bounce point. Land and Furneaux (1997) confirmed this (with more ordinary players). They found that shortly after the opposing player had hit the ball the receiver made a saccade down to a point a few degrees above the bounce point, anticipating the bounce by about 0.2 s (Fig. 17b). At other times the ball was tracked around the table in a normal non-anticipatory way; tracking was almost always by means of saccades rather than smooth pursuit. The reason why players anticipate the bounce is that the location and timing of the bounce are crucial in the formulation of the return shot. Up until the bounce the trajectory of the ball as seen by the receiver is ambiguous. Seen monocularly, the same retinal pattern in space and time would arise from a fast ball on a long trajectory or a slow ball on a short one (Fig. 17a). (Whether either stereopsis or looming information is fast enough to contribute a useful depth signal is still a matter of debate). This ambiguity is removed the instant the timing and position of the bounce are established. Therefore the strategy of the player is to get gaze close to the bounce point (this cannot and need not be exact) before the ball does, and lie in wait. The saccade that affects this is interesting in that it is not driven by a 'stimulus', but by the player's estimate of the location of something that has yet to happen.

2.4.2. *Cricket*

In cricket, where the ball also bounces before reaching the batsman, Land and McLeod (2000) found much the same thing as in table tennis. With fast balls the batsmen watched the delivery and then made a saccade down to the

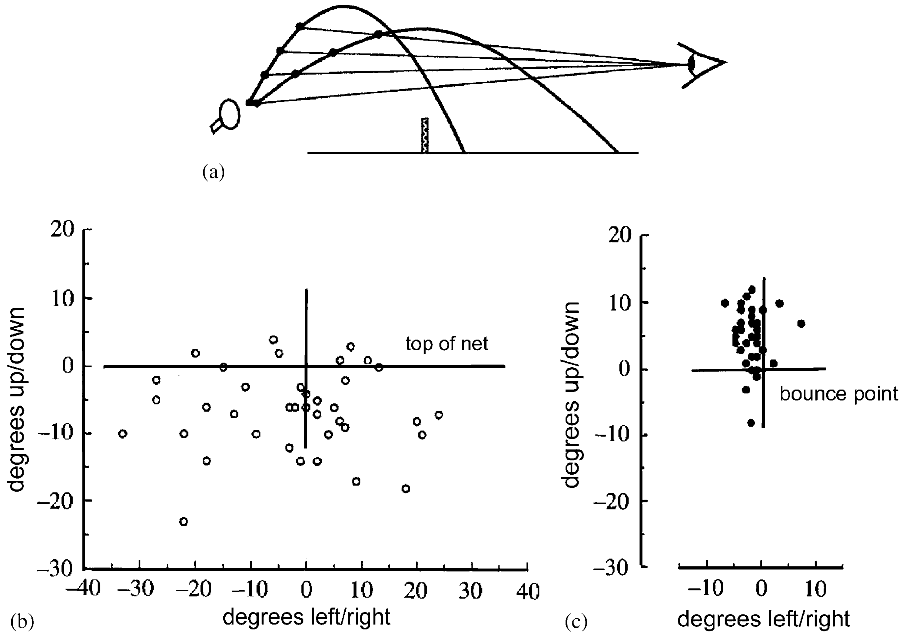


Fig. 17. (a) The visual ambiguity in the trajectory of an approaching ball before it bounces. The vertical motion of a slow ball bouncing short, and a faster ball bouncing long will appear similar to an observer. The ambiguity is removed when the ball bounces. (b,c) The locations in the field of view of the receiver of 38 fixations which follow the first saccade after the ball has been struck by the opponent: (b) relative to the table top, and (c) relative to the bounce point. The receiver mainly fixates a point a few degrees above the expected bounce point, independent of where that is on the table.

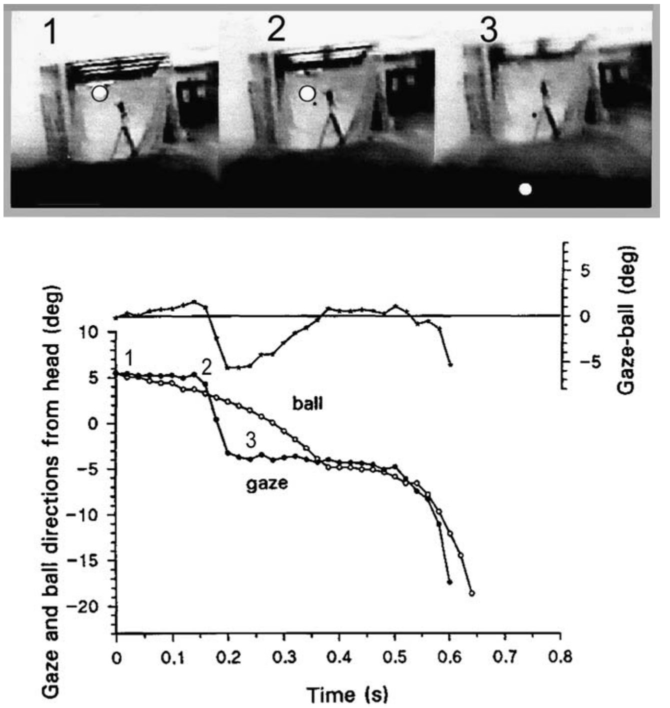


Fig. 18. Upper part: The batsman's view of the ball leaving a bowling machine. (1) Ball about to emerge, batsman's gaze (white dot, 1° across) watching the aperture. (2) Ball (small black dot) descending from the aperture with gaze starting to follow. (3) Gaze saccade to a spot close to the bounce point, which the ball will not reach for a further 0.1 s. Object in centre of each frame is a camera tripod. Lower part, main graph. Vertical direction of gaze (●) and ball (○) viewed from the batsman's head. Numbers correspond to photographs above. Note that the saccade after 2 brings gaze close to the bounce point. After the bounce gaze tracks the ball until about 0.6 s after delivery. The ball is struck at 0.7 s. Upper graph: difference between gaze and eye direction. Note that the batsman must take his eye off the ball by about 5° in order to anticipate the bounce.

bounce point, the eye arriving 0.1 s or more before the ball (Fig. 18). With good batsmen this initial saccade had a latency of only 0.14 s from the time the ball left the bowler's hand, whereas poor or non-batsmen had more typical latencies of 0.2 s or more. This means that poor batsmen can not play really fast balls (~90 mph) because, with the ball taking only 0.4 s to travel the length of the pitch, they are too late to catch the bounce point. Land and McLeod showed that with a knowledge of the time and place of the bounce the batsman has the information he needs to judge where and when the ball will reach his bat, and can thus make an attacking stroke. With slower balls, and balls pitched up so that they bounced closer to the batsman, smooth pursuit was often involved (i.e. the batsmen 'kept the eye on the ball'), but for fast balls pitched short the batsmen always took their eye off the ball by as much as 5° (Fig. 18), the better to see the time, position and behaviour of the ball when it bounced.

2.4.3. Baseball

In baseball the ball does not bounce, and so that source of timing information is not available. Bahill and LaRitz (1984) examined the horizontal head and eye movements to batters facing a simulated fastball. Subjects used smooth pursuit involving both head and eye to track the ball to a point about 9 ft from them, after which the angular motion of the ball became too fast to track (a professional tracked it to 5.5 ft in front: he had exceptional smooth pursuit capabilities). Sometimes batters watched the ball onto the bat by making an anticipatory saccade to the estimated contact point part way through the ball's flight. This may have little immediate value in directing the bat, because the stroke is committed as much as 0.2 s before contact (McLeod, 1987), but may be useful in learning to predict the ball's location when it reaches the bat, especially as the ball often 'breaks' (changes trajectory) shortly before reaching the batter. According to Bahill and LaRitz (1984, p. 253), 'The success of good players is due to faster smooth-pursuit eye movements, a good ability to suppress the VOR, and the occasional use of an anticipatory saccade'.

2.5. Everyday activities involving multiple sub-tasks

Activities such as food preparation, carpentry or gardening typically involve a series of different actions, rather loosely strung together by a 'script'. They provide examples of the use of tools and utensils, and it is of obvious interest to find out how the eyes assist in the performance of these tasks.

2.5.1. Making tea and sandwiches: dividing up the task

Land et al. (1999) studied the eye movements of subjects whilst they made cups of tea. When made with a teapot, this simple task involves about 45 separate acts, acts being defined as 'the movement of an object from one place to another or a change in the state of an object' (Schwartz et al., 1991, p. 384). Fig. 19 shows two examples of the fixations made during the first 10 s of the task. The subjects first examine the kettle, then pick it up and move towards the sink whilst removing the lid from the kettle, place the kettle in the sink and turn on the tap, then watch the water as it fills the kettle. There are impressive similarities both in the form of the scan path, and in the numbers of fixations required for each component of the action (a third subject was also very similar). In each case there is only one fixation that is not directly relevant to the task (the trays to the left of the kettle and to the left of the sink). The two fixations to the right of the sink in JB's record correspond to the place where he put down the lid. Other minor differences concern the timing of the lid removal and details of the way the taps are viewed, but overall the similarities of the two records suggest that the eye movement strategies of different individuals performing similar tasks are highly convergent. The principal conclusions that can be drawn from these scan paths are

(1) Saccades are made almost exclusively to objects involved in the task, even though there are plenty of other objects around to grab the eye.

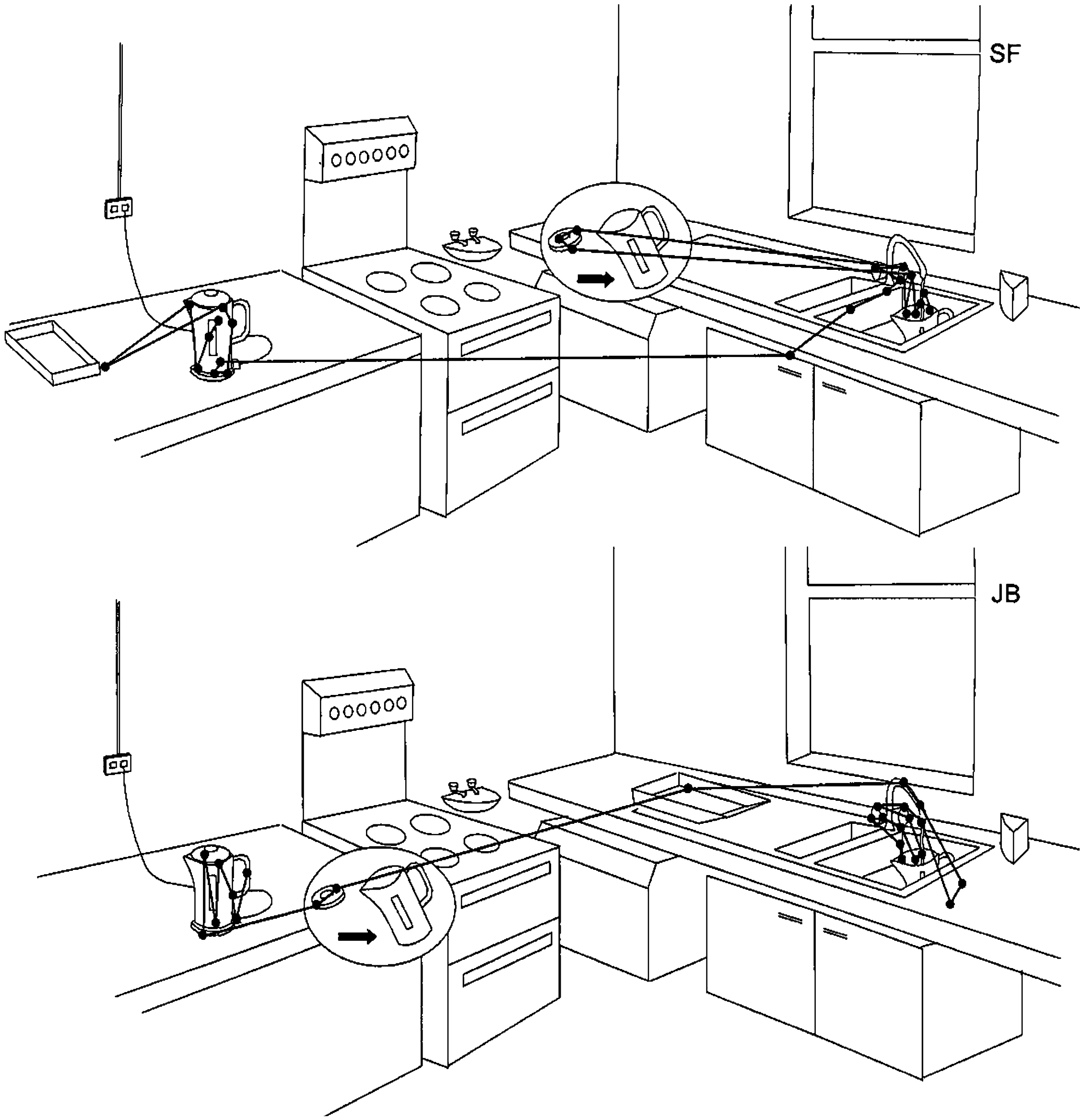


Fig. 19. First 10 s of the tea-making task (filling the kettle) for two subjects. The dots are fixation locations and the lines joining them are saccade trajectories. The two scan paths are similar in the objects visited, the relative numbers of fixations afforded to each object, and the near absence of fixations on 'task-irrelevant' objects.

(2) The eyes deal with one object at a time. This corresponds roughly to the duration of the manipulation of that object, and may involve a number of fixations on different parts of the object.

Almost identical results were obtained by Hayhoe (2000) in a study of students making peanut butter and jelly sandwiches. She found the same attachment of gaze to task-related objects and the same absence of saccades to irrelevant objects. As with the tea-making gaze led manipulation (Fig. 20), although by a somewhat shorter interval. This difference is probably attributable to the fact that the sandwich making was a sit-down task only involving movements of the arms. Two other differences that may have the same cause are the existence of more short duration (<120 ms) fixations than in the tea-making study, and the presence of more 'unguided' reaching movements (13%) mostly concerned with the setting down of objects. There was a clear distinction in both studies between 'within object' saccades which had mean amplitudes of about 8° in both, and 'between object' saccades which were much larger, up to 30° in the sandwich making on a restricted table top, and 90° in tea making in the less restricted kitchen (Land and Hayhoe, 2001).

2.5.2. Timing of movements and actions

In the tea-making study there was usually a clear 'defining moment' when the eyes leave one object and move on to the next, typically with a combined head and eye saccade. These saccades can be used to 'chunk' the task as a whole into separate ‘object-related actions’, and they can act as time markers to relate the eye movements to movements of the body and manipulations by the hands. In this way the different acts in the task can be pooled, to get an idea of the sequence of events in a ‘typical’ act. The results of this are shown in Fig. 20. Perhaps surprisingly, it is the body as a whole that makes the first movement in an object-related action. Often the next object in the sequence is on a different work surface, and this may necessitate a turn or a few steps before it can be viewed and manipulated. About half a second later the first saccade is made to the object, and half a second later still the first indications of manipulation occur. The eyes thus lead the hands. Interestingly, at the end of each action the eyes move on to the next object about half a second before manipulation is complete. Presumably the information that they have supplied remains in a buffer until the motor system requires it.

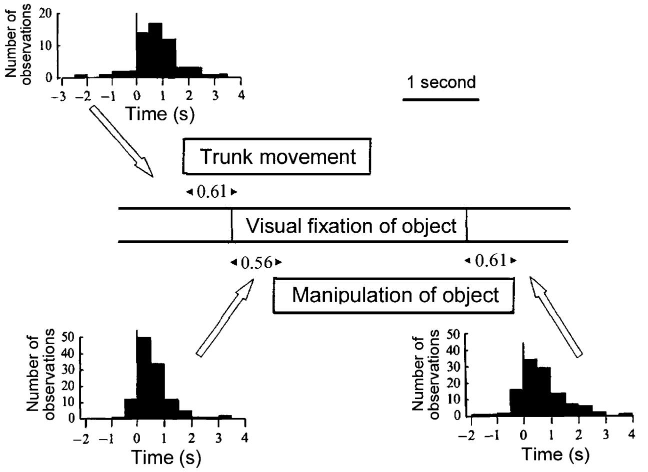


Fig. 20. Average relative timings of trunk movements, fixation of particular objects (which may involve several individual fixations) and manipulation of those objects. Fixations involving long waiting periods (e.g. for the kettle to fill) have been excluded and the data represent 94 object-related actions from three subjects. The typical duration of the visual fixation component is 3.04 s. Trunk movements precede visual fixation, and visual fixations precede the first signs of arm or hand movement each by about 0.6 s. Note that gaze leaves the object about 0.6 s before manipulation is complete, implying that some information is retained in a buffer. From Land et al. (1999).

2.5.3. The functions of single fixations

In tea-making Land et al. (1999) found that about a third of all fixations could be clearly linked to subsequent actions. The remaining two-thirds were made after an action had been initiated so that although they may well have had similar functions in guiding action, it was less clear how they were related to changes in motor behaviour. (Of interest here was a subject, A.I., who had no eye movements, and who made head saccades instead (Land, Furneaux and Gilchrist, 2001; Findlay and Gilchrist, 2003). These are slow and ‘expensive’, and in executing the task shown in Fig. 2 she made only a third as many of these head saccades as normal subjects made eye saccades. As A.I. executed the task with normal speed and competence, the implication is that the rest of us make more saccadic eye movements than we really need.)

Land et al. (1999) found that the functions of fixations could be classified into just four categories, which were designated as locating, directing, guiding and checking (Fig. 21). *Locating* fixations are concerned with establishing the locations of objects, even though there was no associated motor activity at the time of the fixation. In a hand-washing task Pelz and Canoza (2001) also found a number of ‘look ahead’ fixations which anticipated future actions, and these are consistent with this category. *Directing* fixations

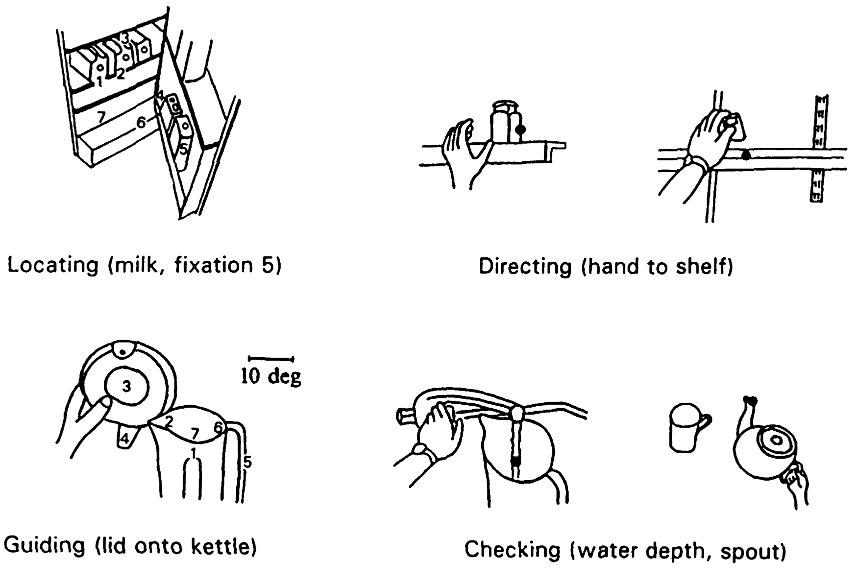


Fig. 21. Four functions of individual fixations identified during the tea-making task. Locating (finding an object, in this case the milk, for future use). Directing (targeting an object or location for immediate action). Guiding (gaze patterns involving manipulations of two or more objects and requiring a number of fixations for completion). Checking (ascertaining whether an action is complete or proceeding properly). Modified from Land et al. (1999).

accompany either a movement of the hand to contact an object (in which case the object is fixated) or the movement of hand and object to a particular position (when the set-down position is fixated). Typically, only a single fixation is involved, and the eye usually moves away from the object or set-down point just before the hand reaches it. Thus visual feedback is not involved in the last stage of the act. It seems that the main function of the directing fixation is to provide fovea-centred goal-position information to guide the arm. *Guiding* fixations are concerned with manipulations involving more than one object, for example a kettle and its lid. Here both objects have to be guided relative to each other so that they dock in an appropriate way. Most tool use is of this nature (e.g. spanner and nut, hammer and nail). It is more complicated than simple directing, and usually involves a number of fixations that alternate between the two objects, and the action is normally completed under visual control. Some guided actions may involve more than two objects, for example knife, bread and butter. *Checking* fixations determine when some condition is met, for example the kettle is full, the water boils, the top is off the bottle. These checking operations may require the eye to dwell on some appropriate region of an object, either in one long fixation or in a series of repeated fixations. When the specified condition is met, a new action is triggered. For example when the kettle is full the tap is turned off. Interestingly, there are some things that are rarely if ever fixated during sequences of this kind. The hands themselves are never fixated, and once objects that have been acquired by the hands they are not fixated either. The implication of this surprising finding is that vision is a scarce and valuable resource, and is disengaged from action as soon as another sense can take over from it.

It is clear that this classification, although appropriate here, does not fit all types of information gathering for which the eyes are employed. In particular a fixation may establish a particular attribute of an object. Thus in the block-copying study of Ballard et al. (1992), shown in Fig. 3 (Section 1.2), the logic of the task dictates that the first fixation on the model establishes the colour of a block, which is then used to locate a similar block for use in the copy. The next fixation on the model is used for another purpose, to check the block's location. In sketching and other copying tasks one or more fixations acquire the form of a feature so that it can be reproduced. Continuous tasks such as reading text or music are also rather different, as each fixation acquires a small part of an information stream which is then compounded internally with those preceding and succeeding it to produce a sequence that has meaning.

3. Issues and conclusions

3.1. Coordination of eye movements and actions

It is clear from the examples given in Part 2 that eye movements play a crucial role in the organization of actions, and that in general the eyes begin to collect information before the action itself has begun (Fig. 20). Eye movements are thus a planned-in, proactive, part of every action sequence, and are not simply summoned up when more information is needed. In the discussions that follow I explore how close are the relations, in space and time, between where we look and what we do, and what is the nature of the control structure that underlies both aspects of visually coordinated activity.

3.1.1. Finding the right information

In driving, ball games and many everyday activities the main function of vision is to supply the motor system with the various kinds of information it needs to carry out the task it is engaged in (see Section 2.5.3 above). In most cases this involves directing the foveas to some specific region of the environment which, at the time, contains the most useful information. Thus in driving it may be the tangent point one moment and a road sign the next, in cricket it may be the bowler's arm or the bounce point of the ball, and in food preparation it may be the handle of the kettle or the carton of milk. Clearly, before it can help the motor system, the visual system needs to know what to look for, where, and when. We have seen that, for familiar tasks, that there is very little uncertainty about this: the eyes go to the appropriate object, in the right place, usually during the second before the associated action begins. Finding the next object in the action sequence is far from simple (see Fig. 22). The object must be recognized either when it is in the visual periphery, or after the fovea has located it using position memory. In either case an appropriate 'search image' must be activated. Positional information can come from two sources: memory and immediate vision. In many instances, during tea making, a subject would make an appropriate turn to the location of an object which was out of sight on the other side of the room, and it is a good assumption that the object's position was established in an earlier phase of the task (in a 'locating' or 'look ahead' fixation). There were several instances during tea-making where an object was fixated a minute or more prior to its ultimate use, and when it was time to fetch the object a 'memory' saccade took gaze to within 10–20° of it. This was then followed by one or two visually guided saccades finally resulting in fixation (Land et al., 1999). Object acquisition thus involves a complex interplay between object recognition, spatial memory and direct vision.

Fig. 22 is an attempt to put into a flow-chart the various operations that must occur during the performance of an 'object-related action' (such as filling a kettle; Fig. 19). We assume that the 'script' of the task provides a succession of low-level 'schemas' which define successive actions (Norman and Shallice, 1986). Each schema specifies the object to be dealt with, the action to be performed on it, and the monitoring required to establish that the action has been satisfactorily completed. First, the gaze control system is supplied with information about the identity and location of the object, the system locates the object, as just described, and then directs the hands to it. The motor

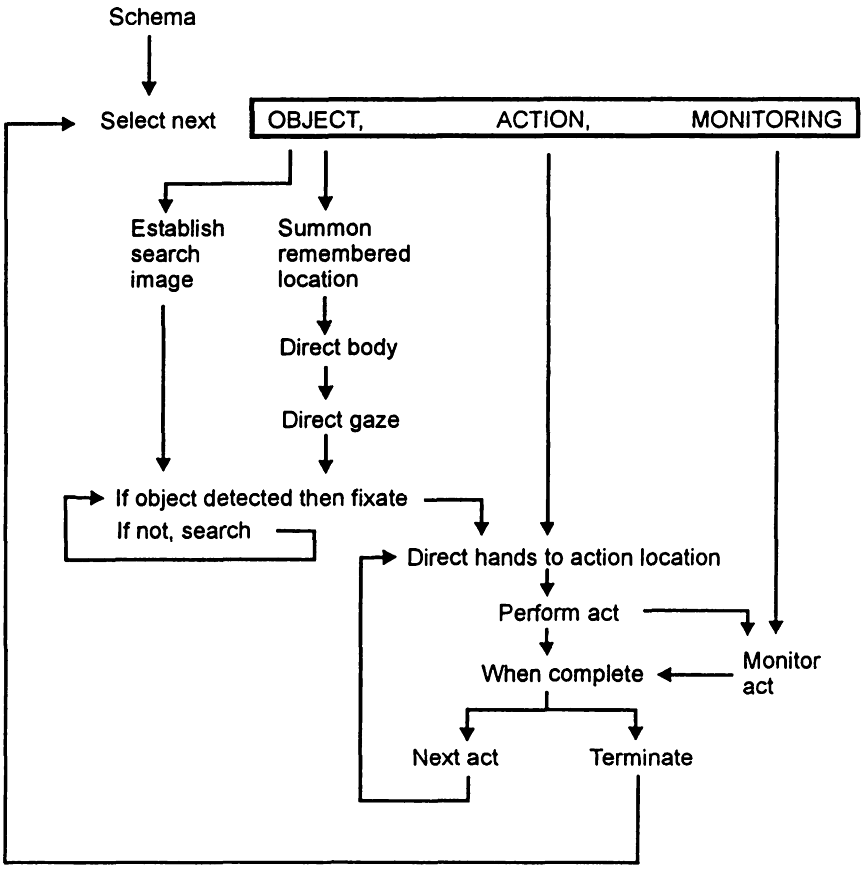


Fig. 22. Diagram indicating the flow of visual information to motor and gaze control systems during the performance of an object-related action. Explanation in text. From Land and Hayhoe (2001).

system of the arms and hands then requires appropriate instructions about the action they must take. Vision may have a simple directing role here, or in the case of 'guiding' (Fig. 21) it may have a more active one, with gaze alternating between two objects as they approach each other. In addition there are checking operations, and the eyes need to know where and what to check. This too must be part of the instructions that the gaze control system receives, presumably at the initiation of each new action. When each sub-act is complete a further manipulation of the same object may follow (lift milk jug, pour milk into cup, set down jug), or if there are no further sub-acts the action as a whole is terminated, the present set of instructions is abandoned and those for the next action are sought.

An interesting question concerns the level of attention and awareness that is involved in the monitoring of object-related actions. In the influential model of Norman and Shallice (1986) the low-level schemas responsible for the control of object-related actions would be considered to be 'routine programs for the control of overlearned skills', not requiring the involvement of conscious attention which is the province of the higher level 'supervisory attention system' called upon during non-routine behaviours, or when problems arise within otherwise routine activities. Nevertheless, we see a surprising amount of monitoring of the routine tasks in food preparation tasks, of which are not conscious, but which is attentive in the sense that it is engaged with the requirements of the task. It seems that, here at least, the connexion between attention and consciousness is not obligatory.

3.1.2. Spatial accuracy: saccade size and scaling

In all the studies given in the first section the eyes look at points that are particularly informative for the ongoing action: in food preparation it is the object being manipulated at the time, in steering it is the tangent point, in ball games the bounce point, etc. How accurately these points are targeted by gaze depends on the spatial scale of the task. Thus in reading each fixation takes in about seven letter spaces, which with standard print at 40 cm means that saccades are about 1.33° long, and so the maximum angular distance of the foveal centre from any one detected letter is half this, 0.67°. At the other extreme, the average size of 'within object' saccades in both tea-making and sandwich making was about 8°, implying a maximum eccentricity of a viewed target of 4° (Land and Hayhoe, 2001), which is 6 times greater than in reading. The difference is presumably due to the large size of culinary objects compared with letters.

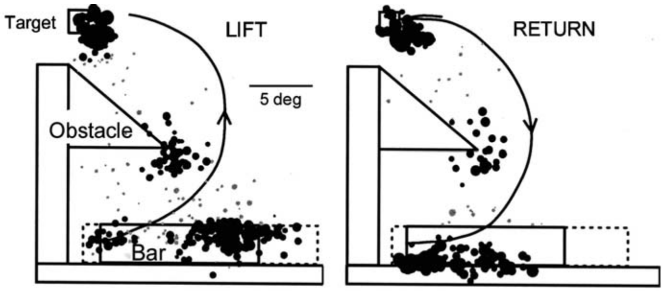


Fig. 23. Accuracy of fixation during a task in which a bar is lifted past an obstacle to make contact with a target and then set down again. All fixations of nine subjects. Five landmarks are consistently viewed. On the LIFT these are the grasp site, the left tip of the bar, the target and the obstacle. On the RETURN the support surface is viewed rather than the grasp site. Black dots are fixations within 3° of one of these landmarks (more than 90%), grey dots are fixations outside these regions. Area of each dot indicates fixation duration. Modified from Johansson et al. (2001).

Most other estimates are between these extreme values. For example Johansson et al. (2001) used a high-precision eye tracker to study performance in a task in which a bar was grasped, and lifted to make contact with a target switch, avoiding a projecting obstacle on the way (Fig. 23). They found that—as in other tasks—the eyes always fixated certain distinct landmarks (the grasp site, the target, and the surface to which the bar returned) and frequently but not always fixated the obstacle and the tip of the bar. They estimated the precision of fixation by determining the sizes of circles that enclosed 90% of the fixation points for all nine subjects: these were 3.3° for the tip of the bar and 5.6° for the obstacle. For the target itself most fixations were within a 3° circle, and they regard 3° as the diameter of the 'functional fovea' for this task. This implies a maximum target eccentricity of 1.5°.

How far from the fovea can useful information be obtained? Again this is likely to depend on the scale of the task, Shioiri and Ikeda (1989) studied the extraction of information from pictures, using a window which was contingent on eye-position. They found that the maximum area over which high-resolution pictorial information could be extracted was about 10° across: larger windows provided no extra information. This implies that no further 'useful resolution' is available outside about 5° from the fixation point. However, this cannot be universally true. Land et al. (1999) found that subjects could make accurate single eye-and-head saccades to appropriate objects that were up to 50° from the current fixation point. Even if position memory is involved here, some information that permits object identification must be available even in the far periphery.

3.1.3. Timing of eye movements and actions

In continuous activities such as reading aloud, sight-reading music and copy-typing we have seen that there is an eye-action delay of about a second. Much of this is no doubt the time needed for the complex transformations that convert written symbols into the actions of muscles. In steering a car there is also a delay of slightly less than a second between viewing tangent point position and adjusting the steering. Besides allowing processing to occur, a second function of the buffer—at least in driving—is that it makes it possible for information for one action to be 'put on hold' whilst the eye deals with something else. A particularly clear instance of this is the behaviour of the motorist to the cyclist in Fig. 13.

In 'one-off' actions such as those involved in food preparation there is also evidence of a buffer in which information for action is stored. Fig. 20 shows that at the beginning of a typical tea-making action the eyes lead manipulative action by about half a second, and at the end they move on to the next object about half a second before the present action is complete. In that half second the final manipulative acts must be guided by information held in a store. It seems likely that throughout each object-related action visual information passes through a half-second buffer, much as in the continuous actions just discussed.

3.2. Conspicuity, instructions and salience

In both tea-making and sandwich-making (Land and Hayhoe, 2001) we were particularly impressed by the way gaze moved from one task-relevant object to the next, ignoring all other objects that were not involved in the activity. The proportion of task-irrelevant objects viewed (other than during periods of waiting—for the kettle to boil for example) was under 5% in both studies. The conclusion must be that—in real tasks—the eyes are driven much more by top-down information from the script of the activity, rather than by the 'intrinsic salience' of objects in the scene. In one sandwich-making experiment involving four subjects, 50% of the objects on the table were irrelevant to the task (pliers, scotch tape, forks, etc.). In the interval before the task commenced, while the eyes were scanning the table, the proportion of irrelevant objects fixated was 52%. When the task started this reduced to 18%. Presumably this represented a shift from target selection based on intrinsic salience to one based on task instructions. Shinoda et al. (2001) reached similar conclusions with a driving task that required the detection of Stop signs; they found that detection was heavily affected by both task instructions and by local context (they were rarely detected mid-block, for example compared with at intersections).

Most recent ideas on the generation of saccades to new targets involve a 'salience map'. This is a two-dimensional surface, tentatively located in the superior colliculus, in which peaks of excitation correspond to objects in the image. These peaks compete with each other in a winner-takes-all manner to reach a threshold that triggers a saccade. Some versions of this salience map concentrate exclusively on bottom-up properties of the image, such as orientation, intensity and colour (e.g. Itti and Koch, 2000), whilst others allow a degree of top-down control to influence the state of the map (Findlay and Walker, 1999). Of particular interest is a study by Torralba (2003) who showed that the task of locating an object in a scene is greatly facilitated if context is taken into account, that is the overall statistical structure of the entire scene. Torralba also showed that context can be used to drive the focus of attention (and so presumably fixations) to regions of the scene where particular objects are likely to be located. Task requirements, however, were not specifically incorporated in this scheme. The studies reported in this review emphasize this latter influence, since it is clear that eye movements are very closely coupled to the script of the action as a whole. As Johansson et al. (2001) put it: 'The salience of potential gaze targets was largely determined by the demands of the sensorimotor task' (p. 6931). The only problem here is that the word 'salience' rather loses its original meaning of conspicuity, and its definition becomes almost circular—an object is salient if it gets looked at it, for whatever reason.

It is not really a surprise to find a strong top-down influence on eye movements in real tasks. In Yarbus' famous recording of eye movements when viewing a picture (Fig. 5) in which observers were asked different questions about the picture, the influence of the instructions in determining the pattern of eye movements was very striking (Yarbus, 1967). In real tasks the instructions come not from the experimenter, but from the learned internal script for the activity. However, the effect in either case is that a large measure of control is imposed over the choice of the objects that are fixated.

3.3. *Roles of different types of memory*

Different types of memory are involved in the interactions between eye movements and actions. First, both the actions themselves, and the associated patterns of eye movements, are achieved by the somewhat mysterious process of procedural learning. This term is applied particularly to the unconscious acquisition of skills (for example learning to ride a bike) by a process that does not involve explicit teaching, and which is not easily verbalized. It contrasts with declarative memory which is brought about by explicit training or stimulus-response pairing. Despite the informal nature of its acquisition, procedural memory can be very durable.

Second, as we have seen in Section 3.1.1, spatial memory for the layout of a room and the position of objects within it is vital if the eyes are to find the objects needed for a task in a quick and efficient manner. The observations that, in a task like food preparation, objects are frequently located by a single saccade and search behaviour is rarely seen, are testimony to the effectiveness of the spatial memory systems involved. The familiarity for overall layout that comes from spending many minutes in a room like a kitchen certainly lasts for days afterwards, and can presumably be considered to be long-term memory. However, the locations of particular moveable objects, such as mugs and kettles, have to be re-acquired each session, and indeed their locations may have to be updated as a session proceeds. Short-term memory lasting from seconds to minutes is presumably involved, and indeed its limited span is entirely appropriate to the task. The 'locating' fixations made during tea-making (Land et al., 1999), and the 'look ahead' fixations described by Pelz and Canoza (2001) during hand washing, are presumably occasions when the locations of particular objects are acquired for future use—notwithstanding the *just in time* strategy that applies much of the time (Ballard et al., 1995; see Section 1.2).

The phenomena of change blindness have demonstrated that a great deal of the information in the image is lost each time we move our eyes (Grimes, 1996; Rensink et al., 2000), but it has also become clear that certain aspects of the scene are retained, although not in the 'pictorial' form of the original image (Henderson and Castelhano, 2005; Tatler et al., 2005). The requirement for many of the tasks discussed here, particularly in Section 2.5, is that the identities and locations of key objects should be retained across many fixations. A possible candidate for the memory system involved here is the 'visuo-spatial sketchpad', a component of the working memory system concerned with tasks that are more visual than verbal (Baddeley, 1997). Although originally devised to deal the manipulation of visual imagery, Baddeley comments (p. 82) that 'It seems likely that the spatial system is important for geographical orientation, and for planning visual tasks'. The capabilities of such a system certainly need to be explored further.

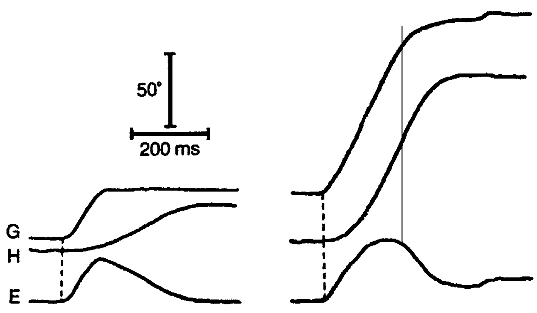
One of the most intriguing forms of memory—if that is the appropriate word—concerns the period between seeing and doing: the period of the eye–voice span in reading aloud, the eye–hand span in typing and music reading, or eye–foot span in walking. This period has a typical duration of slightly less than a second. Information is acquired by the eyes, transformed in various ways, and then emitted as a motor action. The process is continuous, with new information fed onto the conveyor belt as fast as the used information is discarded. It can be interrupted, or 'put on hold' for short periods as in the example of the cyclist and driver in Fig. 13, or when a pianist looks down at the keys. It is thus memory-like in that information is retained, but is not like conventional memory in that it leaves little or no trace; old information is immediately overwritten, it does not just fade away. This process certainly involves a form of working memory, but none of the existing models quite fits. The articulatory loop component of Baddeley's working memory scheme comes close (Baddeley, 1997). In this, a short-term store holds speech sounds for up to 2 s, during which a control process of sub-vocal rehearsal occurs which refreshes items in the store. This model contains two of the elements required for the sort of system needed in other visuo-motor actions: a store and a control process. However, the articulatory loop as proposed by Baddeley is quite specifically geared to the comprehension of speech or the written word, and in particular the 'rehearsal' component that it entails does not seem appropriate to activities such as driving or walking. It may be that there are many loops of this general nature, each with somewhat different properties.

3.4. *Coordination of eyes, head and body*

In natural behaviour the relocations of gaze that are crucial for our visual interactions with the world can be made with eye movements alone, but more commonly they involve combined eye and head movements, or even movements of the eyes, head and trunk. In such vigorous activities as playing tennis it is often necessary to look in one direction whilst rotating the trunk in another, and so there can be no simple rule of thumb that specifies the rotations of the different moveable parts in a gaze rotation of a particular size. In this Section I will consider how such flexibility is achieved.

What combination of eye head and trunk movements occurs depends in part on the size of the gaze rotation required. In the horizontal plane the eyes can rotate up to a maximum of about ±50°, although head movements typically accompany them for gaze rotations larger than 10°. The neck can turn the head through about ±90°. Thus, for gaze rotations greater than 140° the trunk has to become involved, and with freely moving subjects, especially during locomotion, trunk turns can occur with gaze rotations of any amplitude. The eyes rotate faster than the head, which in turn rotates faster than the body. This means that there are potential coordination problems: for example if the eyes reach their target before the head rotation is complete they must counter-rotate as the head catches up, and mechanisms are required to ensure that this can occur without gaze direction being perturbed. Similar considerations apply when part of the rotation is carried out by trunk rotation.

In the case of eye and head rotations of 10° or less Morasso et al. (1973) showed that the trajectory of gaze is the sum of the saccadic eye movements that would have been made with the head still, and a contribution from the semi-circular canals that is the opposite of the rotation made by the head (the VOR). In this way the extent and time course of the gaze saccade is the same, whatever the head does (Fig. 24). However, for saccades beyond the eye movement range (±50°) this will not work, because the eye cannot make the whole saccade on its own. Guitton and Volle (1987) showed that during these large saccades VOR is suspended, so that at least part of the gaze saccade results entirely from rotation of the head. The end-point of the saccade (i.e. VOR is switched on again) is reached when a pre-specified amount of gaze rotation is achieved, as determined by internal signals that monitor head rotation (particular the semi-circular canals, but possibly also neck proprioceptors or efference copy) and eye rotation (efference copy or eye-muscle proprioceptors). Their sum gives gaze rotation (Guitton et al., 2003).



(a) (b)

Fig. 24. Eye–head coordination during saccadic gaze movements. (a) For gaze saccades (G) whose magnitude is well within the eye movement range of about ±50° the gaze change is made mainly with an eye movement (E). The vestibulo-ocular reflex (VOR) ensures that as the head catches up the sum of eye and head movement (gaze movement G) remains constant, and fixation stable. (b) For gaze saccades larger than the eye movement range the onset of VOR is delayed, so that the head carries gaze to the saccade end point, while the eye remains at a plateau position. When VOR is restored (vertical line) gaze direction becomes constant, as in (a)). Note the small 'catch-up' saccade 200 ms later. Modified from Guitton (1992).

What happens when trunk rotation is also involved? A typical situation is illustrated in Fig. 25a. Eye head and trunk start to rotate more or less together (1 and 2), then when gaze reaches the target VOR is switched on counter-rotating the eye and thus maintaining fixation (3 and 4) just as described above. During the last phase (4 and 5) the target is in line with both eye and head, but the body continues to rotate. This means that the head must counter-rotate on the trunk until the trunk too reaches a direction more in line with the target. A recording of a real turn like that in Fig. 25a is shown in Fig. 25b. It shows that the eye behaves as outlined above (Fig. 24b): the saccade is made between 0.4 and 0.6 s, there is a period up to 0.8 s without further motion, and then VOR recommences, with the eye returning to the primary position by 1.5 s. The neck (head on trunk) rotates between 0.4 and 1.0 s, then stops rotating until 1.5 s, but the head continues to rotate in space up to 1.3 s, carried by the trunk. Thereafter the neck counter-rotates as the trunk completes the turn, the trajectory being a near mirror image of the trunk rotation, and the net result is that the head rotates hardly at all in space. Land (2004) has argued that the mechanism that coordinates the counter-rotation of the neck is another reflex, the vestibulo-collic reflex (VCR; see Outerbridge and Melville Jones, 1971; Guitton et al., 1986). Like the VOR this is driven by the semi-circular canals, but differs from the VOR (which is an 'open' feed-forward system) in that it is a feedback loop in which the semi-circular canals provide the basis for an error signal (desired head rotation—zero in this case—minus actual head rotation) which is fed, amplified, to the neck muscles. The result in Fig. 25b is that head rotation in space is 'clamped' close to zero. Head rotation need not necessarily be zero, however. A command to rotate the head can be injected into the input to the loop, and this will be obeyed just like a zero rotation

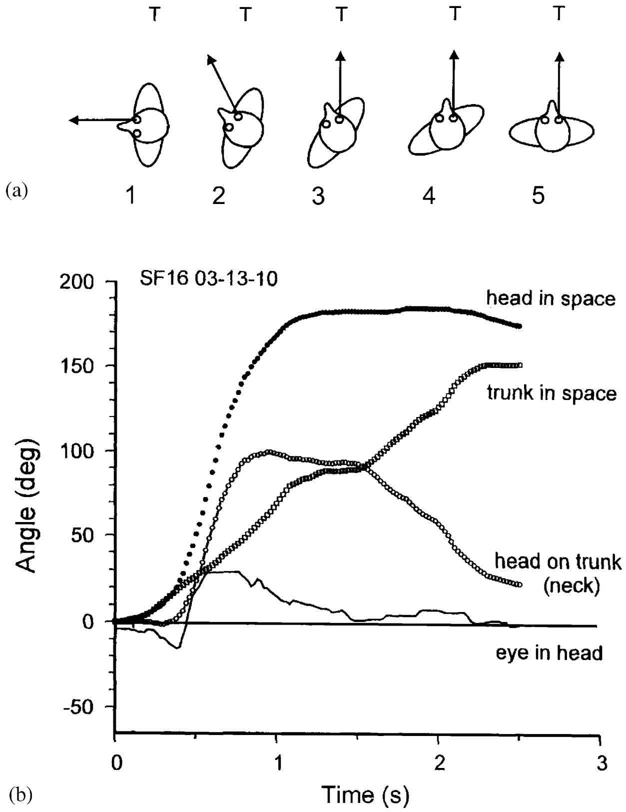


Fig. 25. Coordination of eye, head and trunk rotation during a large saccadic gaze shift. (a) Diagram of a typical 90° turn to a target at T. Eye movement complete by 2; head movement complete by 4, with eyes counter-rotating in head between 2 and 4; trunk continues to rotate until 5, with head counter-rotating on trunk between 4 and 5. (b) Record of a 180° saccadic turn. The eye movement occurs between 0.4 and 0.5 s, after which the eye stops moving in the head until 0.8 s, when the vestibulo-ocular reflex is turned on, and this brings the eye back to the primary position (with three small secondary saccades on the way) as the head continues to rotate. After 1 s the neck begins to counter-rotate as the trunk continues to turn (more obviously after 1.6 s) so that the direction of the head in space is held almost constant. This phase is attributable to the vestibulo-collic reflex. From Land (2004).

command. The system is analogous to power steering in a car.

The combined effect of the two reflexes is that gaze direction can be maintained by VOR independent of rotations of the head, and head direction can be maintained by VCR independent of rotations of the trunk. Without the reflexes, head movements would mechanically cause gaze movements, and trunk movements would mechanically cause head movements. The reflexes permit the controlled emancipation of the different components from the actions of the others.

3.5. *Learning eye–hand coordination*

It has been pointed out repeatedly that each type of action has its own associated regime of eye movements, and that fixations on objects tend to precede actions upon them by up to a second. This is what we observe once an action has been learned, but does this pattern hold during the period while the skill is being acquired? A recent study by Sailer et al. (2005), the first of its kind, has clearly demonstrated that the answer is no.

They devised a task that involved learning how to use a novel mouse-like tool to control the position of a cursor on a screen, in order to hit a series of consecutively appearing target boxes. The tool consisted of a freely held box with a handle at each end. Applying opposite rotational torque to the two handles moved the cursor up and down the screen, whilst pushing the handles towards or away from each other moved the cursor laterally (in fact the system was isometric, and the handles did not move). Making oblique cursor movements with the tool was evidently quite difficult and took some time to learn. The gaze movements and cursor movements of the subjects were monitored as they learned how to use the tool, and measures of success such as hit rate and path length to target were also measured. The learning process took a total of about 20 min.

The most interesting result was that, for most subjects, learning proceeded in three distinct stages: an exploratory phase in which the hit rate remained low, a skill acquisition stage in which the hit rate increased rapidly, and a skill refinement stage in which the rate increased more slowly. The three phases were characterized by very different patterns of both motor control (as shown by the cursor movements) and gaze movements. During the exploratory phase most cursor movements and gaze movements were either horizontal or vertical, as the subjects learned to cope with one or other control dimension of the tool (Fig. 26a). Gaze generally followed the cursor, with occasional glances to the target, and gaze saccades were generally small, 3–4°. At this stage it typically took about 20 s for the cursor to reach the target. During the skill acquisition stage the subjects slowly learned to move the cursor obliquely, and the hit rate increased to about one target every 2 s (Fig. 26b). At the beginning of this second period the eyes continued to track the cursor, although the pattern changed from gaze lagging behind the cursor to leading it by up to 0.3 s: gaze thus began to anticipate cursor position. At the same time gaze saccades became larger, increasing from about 4° to 12°, as more were directed towards the target (successive targets were 18° apart). During the skill refinement stage, gaze no longer tracked the cursor, but went straight to the next target, with either a single saccade, or with one large and one smaller saccade (Fig. 26c). Hit rates increased slightly to just below 1 per second.

The role of gaze clearly changes as learning progresses. To begin with the eyes follow the movements of the cursor. At this stage vision is being used to check on the effects of the rather inexpert control operations, and learning proceeds by associating particular manipulations with their visual consequences. During the skill refinement stage gaze begins to anticipate cursor movements and, as manipulation becomes more secure, vision is used to provide a series

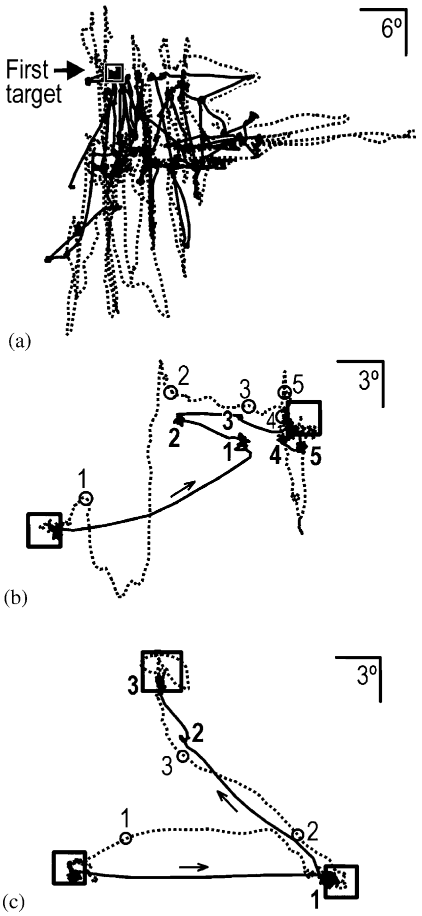


Fig. 26. Three phases of learning a new visuo-motor skill. The task is to use a novel (and rather difficult) mouse-like tool to direct a cursor on a screen to contact targets (boxes) that appear one by one. In the recording dotted lines show the cursor position and full lines the gaze position. Sample records from (a) the exploratory phase, (b) the skill acquisition phase, (c) the skill refinement phase. Further details in text. From Sailer et al. (2005). Copyright by the Society for Neuroscience.

of local goals for cursor movements. Finally, with control of cursor direction established, it is sufficient for gaze to mark the target: i.e. the end point of the cursor movement.

The description of the evolution of eye–hand coordination in this particular skill, with a gradual transition from a monitoring to an anticipatory role for the eyes, sits well with ideas about skill acquisition in other contexts. Many activities we learn after early childhood have components that require the visual calibration of a manual activity—how much to turn the steering wheel of a car, or the effects of a tennis stroke on a ball's trajectory. These imply a period when vision is mainly concerned with checking consequences, corresponding to the first two phases of the scheme proposed by Sailer et al. Equally, once the calibration has been established, vision is freed up to look ahead—to the next bend or the intended target of the next stroke (Fig. 16, of the learner driver, is relevant here). Vision comes to adopt a feed-forward rather than a feedback role.

It has long been recognized that skill acquisition proceeds in stages. Psychologists distinguish the early attention-demanding stages in learning a new skill, from the later stages in which actions are automatized, and require little or no conscious intervention (Norman and Shallice, 1986; Underwood and Everatt, 1996). There are also theoretical models of the internal processes involved in motor control which allow for learning by comparing motor signals with sensory feedback, and which lead to predictive behaviour—the kinds of processes that seem to be at work here (Miall and Wolpert, 1996). The Sailer et al. study has provided important details about the ways that eye movements contribute to the learning process, with which it should be possible to refine models of skill acquisition.

3.6. Future directions

One of the main conclusions from this review is that eye movement strategies are very task-specific. They principally involve the acquisition of information needed for the execution of motor actions in the second or so before they are performed, and in the checking of the execution of each action. Although the work of the last two decades has provided a reasonable sample of the ways that eye movements and actions are related in different types of task, there are many others that might be studied, in domestic and industrial situations, and in other competitive sports. There is much to learn about the way that eye movement patterns are learned. The new work of Sailer et al. (2005), outlined in Section 3.5, has provided a valuable starting point for such studies, but there is much more here to be found out. The question of how eye movement strategies are acquired in childhood has not yet been touched, and even for tasks learned in adulthood we only have a single, rather artificial, example.

An important unanswered question concerns that way that the visual system is able to locate the objects needed in each task, in the right order and at the right time. It is now very clear that the intrinsic salience of objects, based on their bottom-up image statistics, has little to do with this, and that, in tasks like food preparation and driving, objects are located on the basis of their immediate relevance to the task. The nature of the internal filing arrangements that must be brought to bear to achieve this are largely unexplored. They involve place memory, but also a feature identification system that allows the appropriate objects to be targeted and recognized both on and outside the fovea. Judged by the rarity of search behaviour in the performance of everyday tasks, this must be a very efficient system.

The neurophysiology and anatomy of the eye movement and action systems has been outside the scope of this review, but it is probably true to say that whilst a great deal is known about the neural basis of individual eye movements and actions, there is no equivalent body of knowledge relation to eye movement and motor strategies (but see Hayhoe and Ballard, 2005). Since in this review, we have been looking mainly at natural behaviour, and it has never been easy to explore the workings of the brain during unrestrained activity, acquiring this knowledge is likely to be a daunting task. Scanning studies may be helpful, but again not much real activity is possible within a scanner. The neuropsychological investigation of behavioural deficits is likely to be the most productive methodology in the short term.

References

Antes, J.R., 1974. The time course of picture viewing. J. Exp. Psychol. 103, 62–70.

Baddeley, A., 1997. Human Memory, Theory and Practice, second ed. Psychology Press, Hove.

Bahill, A.T., LaRitz, T., 1984. Why can't batters keep their eyes on the ball? Am. Sci. 72, 249–253.

Ballard, D.H., Hayhoe, M.M., Li, F., Whitehead, S.D., 1992. Hand–eye coordination during sequential tasks. Philos. Trans. R. Soc. Lond. B 337, 331–339.

Ballard, D.H., Hayhoe, M.M., Pelz, J.B., 1995. Memory representations in natural tasks. J. Cogn. Neurosci. 7, 66–80.

Buswell, G.T., 1920. An experimental study of the eye-voice span in reading. Supplementary Educational Monographs No. 17.

Buswell, G.T., 1935. How People Look at Pictures: A Study of the Psychology of Perception in Art. Chicago University Press, Chicago.

Butsch, R.L.C., 1932. Eye movements and the eye–hand span in typewriting, J. Educ. Psychol. 23, 104–121.

Carpenter, R.H.S., 1988. Movements of the Eyes. Pion, London.

Carpenter, R.H.S., 1991. The visual origins of ocular motility. In: Carpenter, R.H.S. (Ed.), Vision and Visual Dysfunction. v.8. Eye Movements. Macmillan, Basingstoke, pp. 1–10.

Collewijn, H., 1977. Gaze in freely moving subjects. In: Baker, R., Berthoz, A. (Eds.), The Control of Gaze by Brainstem Neurons. Elsevier, New York, pp. 13–22.

Collewijn, H., Martins, A.J., Steinman, R.M., 1981. Natural retinal image motion: origin and change. Ann. N.Y. Acad. Sci. 374, 312–329.

Crundall, D., 2005. The integration of top-down and bottom-up factors in visual search during driving. In: Underwood, G. (Ed.), Cognitive Processes in Eye Guidance. Oxford University Press, Oxford, pp. 283–302.

Crundall, D., Underwood, G., Chapman., P., 1998. How much do novice drivers see? The effects of demand on visual search strategies in novice and experienced drivers. In: Underwood, G. (Ed.), Eye Guidance in Reading and Scene Perception. Elsevier, Amsterdam, pp. 395–417.

Delabarre, E.B., 1898. A method for recording eye-movements. Am. J. Psychol. 9, 572–574.

Dodge, R., 1900. Visual perception during eye movement. Psychol. Rev. 7, 454–465.

Dodge, R., Cline, T.S., 1901. The angle velocity of eye movements. Psychol. Rev. 8, 145–157.

Donges, E., 1978. A two-level model of driver steering behavior. Hum. Factors 20, 691–707.

Duchowski, A.T., 2003. Eye Tracking Methodology: Theory and Practice. Springer, Berlin.

Erdmann, B., Dodge, R., 1898. Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage. Niemeyer, Halle.

Findlay, J.M., Gilchrist, I.D., 2003. Active Vision. Oxford University Press, Oxford.

Findlay, J.M., Walker, R., 1999. A model of saccade generation based on parallel processing and competitive inhibition. Behav. Brain Sci. 22, 661–721.

Fleischer, A.G., 1986. Control of eye movements by working memory load. Biol. Cybernet. 55, 227–238.

Friedburg, C., Allen, C.P., Mason, P.J., Lamb, T.D., 2004. Contribution of cone photoreceptors and postreceptoral mechanisms to the human photopic electroretinogram. J. Physiol. (Lond.) 556, 819–843.

Furneaux, S., Land, M.F., 1999. The effects of skill on the eye–hand span during musical sight-reading. Proc. R. Soc. Lond. B 266, 2435–2440.

Gandhi, N.J., Sparks, D.L., 2003. Changing views of the role of superior colliculus in the control of gaze. In: Chalupa, L.M., Werner, J.S. (Eds.), The Visual Neurosciences. MIT Press, Cambridge MA, pp. 1449–1465.

Grimes, J., 1996. On the failure to detect changes in scenes across saccades. In: Atkins, K. (Ed.), Perception: Vancouver Studies in Cognitive Science, vol. 2. Oxford University Press, New York, pp. 89–110.

Guitton, D., 1992. Control of eye–head coordination during orienting gaze shifts. Trends Neurosci. 15, 174–179.

Guitton, D., Volle, M., 1987. Gaze control in humans: eye–head coordination during orienting movements to targets within and beyond the oculomotor range. J. Neurophysiol. 58, 427–459.

Guitton, D., Kearney, R.E., Wereley, N., Peterson, B.W., 1986. Visual, vestibular and voluntary contributions to human head stabilization. Exp. Brain Res. 64, 59–69.

Guitton, D., Bergeron, A., Choi, W.Y., Matsuo, S., 2003. On the feedback control of orienting gaze shifts made with eye and head movements. Prog. Brain Res. 142, 55–68.

Hayhoe, M., 2000. Vision using routines: a functional account of vision. Vis. Cogn. 7, 43–64.

Hayhoe, M., Ballard, D., 2005. Eye movements and natural behavior. Trends Cogn. Sci. 9, 188–194.

Henderson, J.M., Castelhano, M.S., 2005. Visual memory for scenes. In: Underwood, G. (Ed.), Cognitive Processes in Eye Guidance. Oxford University Press, Oxford, pp. 213–235.

Henderson, J.M., Hollingworth, A., 1998. Eye movements during scene viewing: an overview. In: Underwood, G. (Ed.), Eye Guidance in Reading and Scene Perception. Elsevier, Amsterdam, pp. 269–293.

Hollands, M.A., Patla, A.E., Vickers, J.N., 2002. "Look where you're going!": gaze behaviour associated with maintaining and changing the direction of locomotion. Exp. Brain Res. 143, 221–231.

Imai, T., Moore, S.T., Raphan, T., Cohen, B., 2001. Interaction of the body, head, and eyes during walking and turning. Exp. Brain Res. 136, 1–18.

Inhoff, A.W., Wang, J., 1992. Encoding of text, manual movement planning, and eye–hand coordination during copy-typing. J. Exp. Psychol.: Hum. Percept. Perform. 18, 437–448.

Itti, L., Koch, C., 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Res. 40, 1489–1506.

Johansson, R.S., Westling, G., Bäckström, A., Flanagan, J.R., 2001. eye–hand coordination in object manipulation. J. Neurosci. 21, 6917–6932.

Kowler, E., 1991. The stability of gaze and its implications for vision. In: Carpenter, R.H.S. (Ed.), Vision and Visual Dysfunction, vol. 8. Eye Movements. Macmillan, Basingstoke, pp. 71–92.

Land, M.F., 1993. Eye head co-ordination during driving. in: Proceedings of the IEEE Systems, Man and Cybernetics Conference, Le Touquet, Vol. 3, pp. 490-494.

Land, M.F., 1998. The visual control of steering. In: Harris, L.R., Jenkin, M. (Eds.), Vision and Action. Cambridge University Press, Cambridge, pp. 163–180.

Land, M.F., 1999. Motion and vision: why animals move their eyes. J. Comp. Physiol. A 185, 341–352.

Land, M.F., 2004. The coordination of rotations of eyes, head and trunk in saccadic turns produced in natural situations. Exp. Brain Res. 159, 151–160.

Land, M.F., Furneaux, S., 1997. The knowledge base of the oculomotor system. Philos. Trans. R. Soc. Lond. B 352, 1231–1239.

Land, M.F., Hayhoe, M., 2001. In what ways do eye movements contribute to everyday activities. Vision Res. 41, 3559–3565.

Land, M.F., Horwood, J., 1995. Which parts of the road guide steering? Nature 377, 339–340.

Land, M.F., Lee, D.N., 1994. Where we look when we steer. Nature 369, 742–744.

Land, M.F., McLeod, P., 2000. From eye movements to actions: how batsmen hit the ball. Nat. Neurosci. 3, 1340–1345.

Land, M.F., Tatler, B.W., 2001. Steering with the head: the visual strategy of a racing driver. Curr. Biol. 11, 1215–1220.

Land, M.F., Mennie, N., Rusted, J., 1999. The roles of vision and eye movements in the control of activities of daily living. Perception 28, 1311–1328.

Land, M.F., Furneaux, S.M., Gilchrist, I.D., 2001. The organization of visually mediated actions in a subject without eye movements. Neurocase 8, 80–87.

Mackworth, N.H., Thomas, E.L., 1962. Head-mounted eye-movement camera. J. Opt. Soc. Am. 52, 713–716.

McLeod, P., 1987. Visual reaction time and high-speed ball games. Perception 16, 49–59.

Miall, R.C., Tchalenko, J., 2001. A painter's eye movements: a study of eye and hand movement during portrait drawing. Leonardo: Int. J. Arts Sci. 34, 35–40.

Miall, R.C., Wolpert, D.M., 1996. Forward models for physiological motor control. Neural Networks 9, 1265–1279.

Miura, T., 1987. Behavior oriented vision: functional field of view and processing resources. In: O'Regan, J.K., Lévy-Schoen, A. (Eds.), Eye Movements: From Physiology to Cognition. North-Holland, Amsterdam, pp. 563–572.

Morasso, R., Bizzi, E., Dichgans, J., 1973. Adjustments of saccade characteristics during head movements. Exp. Brain Res. 16, 492–500.

Mourant, R.R., Rockwell, T.H., 1970. Mapping eye-movement patterns to the visual scene in driving: an exploratory study. Hum. Factors 12, 81–87.

Mourant, R.R., Rockwell, T.H., 1972. Strategies of visual search by novice and experienced drivers. Hum. Factors 14, 325–335.

Norman, D.A., Shallice, T., 1986. Attention to action: willed and automatic control of behavior. In: Davidson, R.J., Schwarts, G.E., Shapiro, D. (Eds.), Consciousness and Self Regulation. Advances in Research and Theory, vol. 4. Plenum Press, New York, pp. 1–18.

Outerbridge, J.S., Melville Jones, G., 1971. Reflex vestibular control of head movements in man. Aerospace Med. 42, 935–940.

Patla, A.E., Vickers, J.N., 2003. How far ahead do we look when required to step on specific locations in the travel path during locomotion. Exp. Brain Res. 148, 133–138.

Patla, A.E., Prentice, S.D., Rietdyk, S., Allard, F., Martin, C., 1999. What guides the selection of alternate foot placement during locomotion in humans. Exp. Brain Res. 128, 441–450.

Pelz, J.B., Canoza, R., 2001. Oculomotor behavior and perceptual categories in complex tasks. Vision Res. 41, 3587–3596.

Radach, R., Kennedy, A., Rayner, K. (Eds.), 2004. Eye Movements and Information Processing During Reading. Psychology Press, New York.

Ratliff, F., Riggs, L.A., 1950. Involuntary motions of the eye during monocular fixation. J. Exp. Psychol. 40, 687–701.

Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. Psychol. Bull. 124, 372–422.

Regan, D., 1992. Visual judgments and misjudgments in cricket, and the art of flight. Perception 21, 91–115.

Rensink, R.A., O'Regan, J.K., Clark, J.J., 2000. On the failure to detect changes in scenes across brief interruptions. Vis. Cogn. 7, 127–145.

Ripoll, H., Fleurance, P., Caseneuve, D., 1987. Analysis of visual patterns of table tennis players. In: O'Regan, J.K., Levy-Schoen, A. (Eds.), Eye Movements: From Physiology to Cognition. North-Holland, Amsterdam, pp. 616–617.

Robinson, D.A., 1968. Eye movement control in primates. Science 161, 1219–1224.

Robinson, D.A., 1981. Control of eye movements. In: Brooks, V.B. (Ed.), Handbook of Physiology, Section 1, The Nervous System, vol. II, part 2. American Physiological Society, Bethesda, pp. 1275–1320.

Sailer, U., Flanagan, J.R., Johansson, R.S., 2005. eye–hand coordination during learning of a novel visuomotor task. J. Neurosci. 25, 8833–8842.

Schall, J.D., 2003. Selection of targets for saccadic eye movements. In: Chalupa, L.M., Werner, J.S. (Eds.), The Visual Neurosciences. MIT Press, Cambridge MA, pp. 1369–1390.

Schwartz, M.F., Reed, E.S., Montgomery, M.W., Palmer, C., Mayer, N.H., 1991. The quantitative description of action disorganisation after brain damage: a case study. Cogn. Neuropsychol. 8, 381–414.

Shinoda, H., Hayhoe, M.M., Shrivastava, A., 2001. What controls attention in natural environments? Vision Res. 41, 3535–3545.

Shioiri, S., Ikeda, M., 1989. Useful resolution for picture perception as a function of eccentricity. Perception 18, 347–361.

Sommer, M.A., Wurtz, R.H., 2003. The dialogue between cerebral cortex and superior colliculus: implications for target selection and corollary discharge. In: Chalupa, L.M., Werner, J.S. (Eds.), The Visual Neurosciences. MIT Press, Cambridge MA, pp. 1466–1484.

Steinman, R.M., 2003. Gaze control under natural conditions. In: Chalupa, L.M., Werner, J.S. (Eds.), The Visual Neurosciences. MIT Press, Cambridge MA, pp. 1339–1356.

Summala, H., 1998. Forced peripheral vision driving paradigm: evidence for the hypothesis that car drivers learn to keep in lane with peripheral vision. In: Gale, A.G. (Ed.), Vision in Vehicles VI. Elsevier, Amsterdam, pp. 51–60.

Summala, H., Nieminen, T., Punto, M., 1996. Maintaining lane position with peripheral vision during in-vehicle tasks. Hum. Factors 38, 442–451.

Tatler, B.W., Baddeley, R.J., Gilchrist, I.D., 2005. Visual correlates of fixation selection: effects of scale and time. Vision Res. 45, 643–659.

Tatler, B.W., Gilchrist, I.D., Land, M.F., 2005. Visual memory for objects in natural scenes: from fixations to object files. Quart. J. Exp. Psychol. 58, 931–960.

Tchalenko, J., Dempere-Marco, L., Hu, X.P., Yang, G.Z., 2003. Eye movement and voluntary control in portrait drawing. In: Hyönä, J.R., Deubel, H. (Eds.), The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research. Elsevier, Amsterdam, pp. 705–727.

Thomas, E.L., 1968. Movements of the eye. Sci. Am. 219 (2), 88–95.

Torralba, A., 2003. Contextual priming for object detection. Int. J. Comput. Vision 53, 160–191.

Underwood, G., Everatt, J., 1996. Automatic and controlled information processing: the role of attention in the processing of novelty. In: Neumann, O., Sauders, A.F. (Eds.), Handbook of Perception and Action, vol. 3. Attention. Academic Press, London, pp. 185–227.

Underwood, G., Chapman, P., Crundall, D., Cooper, S., Wallén, R., 1999. The visual control of steering and driving: where do we look when negotiating curves. In: Gale, A.G. (Ed.), Vision in Vehicles VII. Elsevier, Amsterdam, pp. 245–252.

Wade, N.J., Tatler, B.W., 2005. The Moving Tablet of the Eye: the Origins of Modern Eye Movement Research. Oxford University Press, Oxford.

Walls, G.L., 1962. The evolutionary history of eye movements. Vision Res. 2, 69–80.

Weaver, H.E., 1943. A study of visual processes in reading differently constructed musical selections. Psychol. Monogr. 55, 1–30.

Westheimer, G.A., McKee, S., 1975. Visual acuity in the presence of retinal image motion. J. Opt. Soc. Am. 65, 847–850.

Wertheim, T., 1894. Über die indirekte Sehshärfe. Zeitschr. Psychol. Physiol. Sinnesorgane 7, 172–187.

Yarbus, A., 1967. Eye Movements and Vision. Plenum Press, New York.

Yilmaz, E.R., Nakayama, K., 1995. Fluctuation of attention levels during driving, Invest. Ophthalmol. Vis. Sci. 36, S940.

Zwahlen, H.T., 1993. Eye scanning rules for drivers: how do they compare with actual observed eye-scanning behavior? Transport. Res. Rec. 1403, 14–22.