Eye movements in natural behavior

Mary Hayhoe and Dana Ballard

Departments of Brain & Cognitive Science and Computer Science, University of Rochester, Rochester, NY 14627, USA

The classic experiments of Yarbus over 50 years ago revealed that saccadic eye movements reflect cognitive processes. But it is only recently that three separate advances have greatly expanded our understanding of the intricate role of eye movements in cognitive function. The first is the demonstration of the pervasive role of the task in guiding where and when to fixate. The second has been the recognition of the role of internal reward in guiding eye and body movements, revealed especially in neurophysiological studies. The third important advance has been the theoretical developments in the fields of reinforcement learning and graphic simulation. All of these advances are proving crucial for understanding how behavioral programs control the selection of visual information.

Over 50 years ago a Russian scientist, Alfred Yarbus, was able to capture the movements of the eye by attaching a mirror system to the eyeball. Although he was not the first to measure eye movements, his work most clearly called attention to their intrinsically cognitive nature. Nowadays most vision scientists are familiar with his traces of a subject examining Repin's painting: 'They Did Not Expect Him', and the very different gaze patterns elicited by different instructions to the subject [1]. The significance of this finding was that it revealed in a particularly compelling way that 'seeing' is inextricably linked to the observer's cognitive goals. Although the relationship between eye movements and cognitive processes has been studied extensively since that time [2–4], it is only recently that several separate lines of inquiry have coalesced to give an increasingly coherent understanding of the elaborate and intricate role of eye movements in cognitive function. The intent of this review is to describe these advances.

There have been many recent developments in the study of cognition in eye movements that have been comprehensively described in [4–7]. This review focuses on simple natural behaviors and describes three principal complementary advances that have had a direct bearing on their study. The first is the description of the role of eye movements in executing everyday visually guided behaviors [8–12]. These measurements have been driven by the development of portable eye trackers that can be worn by subjects engaged in behaviors that involve substantial body movements [13–15] (Box 1). The major findings of these studies have been the importance of the role of the task in such movements, and in learning where and when

Corresponding authors: Hayhoe, M. (mary@cvs.rochester.edu), Ballard, D. (dana@cs.rochester.edu).

to fixate. The second advance has been the recognition of the role of internal reward in guiding eye and body movements [16–20]. This has been revealed especially in neurophysiological studies. The third important advance has been the theoretical developments in the field of reinforcement learning, together with tremendous advances in graphic simulation [21–23]. Together, these developments have allowed the simulation of reward-based systems that incorporate realistic models of eye movements over extended time scales. This has shifted the

---

Box 1. Eye tracking in natural behavior

Fifty years ago, Yarbus' subjects had to do the experiments with a mirror attached to a suction cup on the sclera of the eye. Subsequent systems, such as eye coils and the Dual-Purkinje Image Tracker were more comfortable, but still required the head to be fixed. However, the past ten years has seen the advent and rapid refinement of portable eye trackers that allow eye tracking during free viewing. Michael Land built one of the first of these but it was hampered by having to have the eye position calculated off line [8–10]. Modern eye trackers have fast software to do this and a recent version by Pelz (Figure I) is completely portable, using power from a backpack-mounted battery pack [15]. The new eye trackers allow the study of eye movements over extended tasks in natural settings, where a much wider variety of natural coordinated behaviors is possible.

Figure I. Portable, head mounted eye-tracker developed by Pelz and colleagues [15], based on commercially available systems that use an infrared video camera to image pupil and the corneal reflection. A camera mounted on the frame of the glasses records the scene from the observer's viewpoint. Eye position is then superimposed on the video record.

focus of experimental understanding from where in a scene the eyes fixate in an image, to why the eyes choose a location in a scene, and when they choose it.

Eye tracking in natural behavior

The importance of task

Although Yarbus' study revealed the importance of the instructions in determining where subjects look, the particular fixations did not reveal much more than that. This problem continues to be particularly acute in situations where subjects passively view pictures where the experimenter often has little control of, and no access to, what the observer is doing, although some regularities in fixation patterns can be explained by image properties such as contrast or chromatic salience [24–26]. By contrast, recent experiments where the task structure is evident have been much more easily interpreted, because the task provides an external referent for the internal computations [2,27,28]. One criticism of the stress on task context is that such effects can be covert, but there are typically many ways in which an experiment can be structured to make attentional shifts overt (e.g. see [29,30]). The most novel finding of task-oriented studies is that the eyes are positioned at a point that is not the most visually salient, but is the best for the spatio-temporal demands of the job that needs to be done. This line of investigation has been used in extended visuo-motor tasks such as driving, walking, sports, and making tea or sand-wiches [8–10,12,31–33]. The central result of all these investigation is that fixations are tightly linked in time to the evolution of the task. Very few irrelevant areas are fixated. Figure 1 shows the clustering of fixations on task-specific regions when a subject makes a sandwich. Figure 2 shows the tight linkage, in time, of the fixations to the actions. Ballard *et al.* [28] called this a 'just-in-time' strategy, where observers acquire the specific information they need just at the point it is required in the task.
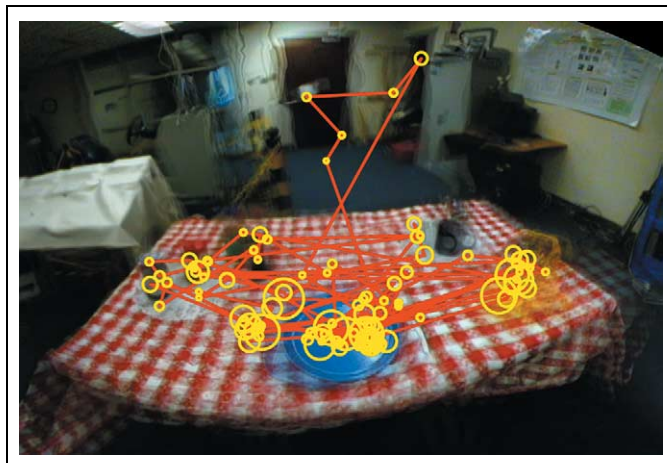
Specialized computations during fixations

Not only is the sequence of fixations tightly linked to the task, but in addition, fixations appear to have the purpose of obtaining quite specific information. For example, cricket players fixate the bounce point of the ball just ahead of its impact, as the location and time of the bounce provide batsmen with the information they need to estimate the desired contact point with the bat [33]. Box 2 provides further evidence that highly task-specific information is extracted in different fixations. These task-specific computations have been referred to as 'visual routines' [36–38]. This specificity is indicated not only by the ongoing actions and the point in the task, but also by the durations of the fixations, which vary over a wide range [11,12]. It appears that a large component of this variation depends on the particular information required for that point in the task, fixation being terminated when the particular information is acquired [4,14,39]. This underscores the overriding control of visual operations by the internal agenda rather than the properties of the stimulus, and the range of different kinds of visual information that can be extracted from the same visual stimulus.

Learning where to look

Implicit in much of the research on natural tasks is the finding that eye movement patterns must be learned [9,13,41]. For example, in tea making and sandwich making (Figure 1), observers must have learnt what objects in the scene are relevant, because almost no fixations fall on irrelevant objects. In driving, Shinoda *et al.* [42] showed that approximately 45% of fixations fell
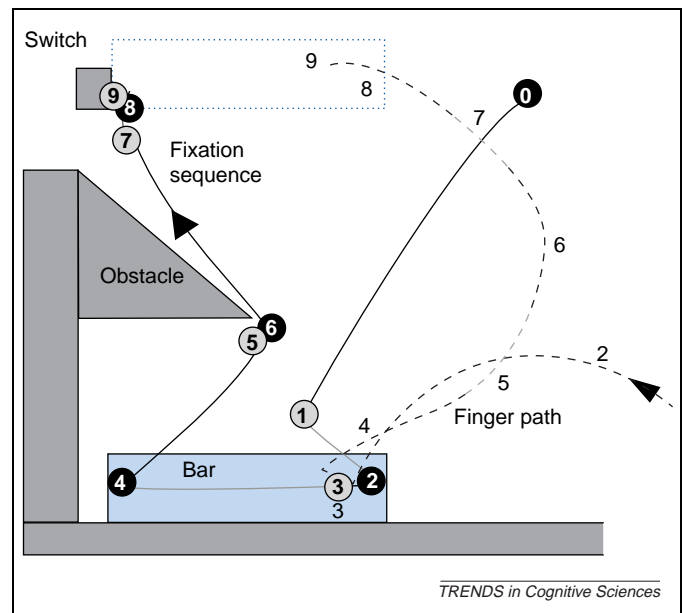

Figure 1. Fixations made by an observer while making a peanut butter and jelly sandwich. Images were taken from a camera mounted on the head, and a composite image mosaic was formed by integrating over different head positions using a method described in Rothkopf and Pelz [34]. (The reconstructed panorama shows artifacts due to the incomplete imaging model that does not take the translational motion of the subject into account.) Fixations are shown as yellow circles, with diameter proportional to fixation duration. Red lines indicate the saccades. Note that almost all fixations fall on task-relevant objects.

Figure 2. Link between fixation and the hand. Fixation locations and hand path are shown for a task in which a subject picks up and moves a bar (blue) past an obstacle, and then contacts a switch. The solid line indicates the fixations, and the dashed line shows the fingertip position. The subject holds the bar by the right end and must maneuver the tip past the triangular shaped obstacle. Numbers on the fingertip path indicate fingertip position during the fixation with the corresponding number. Fixations are made at critical points such as the point of contact for the fingers, then the end of the bar after pickup, followed by fixation near the tip of the obstacle while the bar is moved around it, and then near the switch once the bar has cleared the obstacle. Adapted from Johansson *et al.* [35].

Box 2. What do we see when we look?

Whereas a given cognitive event might reliably lead to a particular fixation, the fixation itself does not uniquely specify the cognitive event. Fixation patterns in natural tasks suggests that very specific visual computations are performed at the fixation point as needed for task performance [8,11,12]. An experiment by Triesch *et al.* [40] supports this supposition. Subjects sorted virtual bricks of two different heights onto two 'conveyor belts' (horizontal strips on the right hand side of the workspace; see Figure I) according to different rules that vary the points at which the brick height is relevant in the task. In one condition, subjects picked up the bricks in front-to-back order and placed them on a belt. In a second condition, subjects picked up the tall bricks first and put them on a belt, and then, picked up the small bricks and put them on the same belt. In a third condition, the tall bricks were put on the front belt, and the short bricks on the back belt.

In the first case, size is irrelevant. In the second, size is relevant for pickup only. In the third, it is relevant for both pickup and placement. On some trials, the brick changed size while it was being moved to the belt. Subjects rarely noticed the change when size was irrelevant, suggesting they did not visually represent the brick size in these trials. Interestingly, when size was relevant only for pickup, subjects were less likely to notice changes than when it was relevant for placement as well, suggesting they did not retain a representation of size in working memory when it was no longer needed. On some trials, subjects tracked the brick while they moved it across the workspace, so that they were fixating the brick at the point when it changed, but were still unaware of the change. This suggests that subjects may not represent particular stimulus features such as size when they are performing other computations such as guiding the arm.
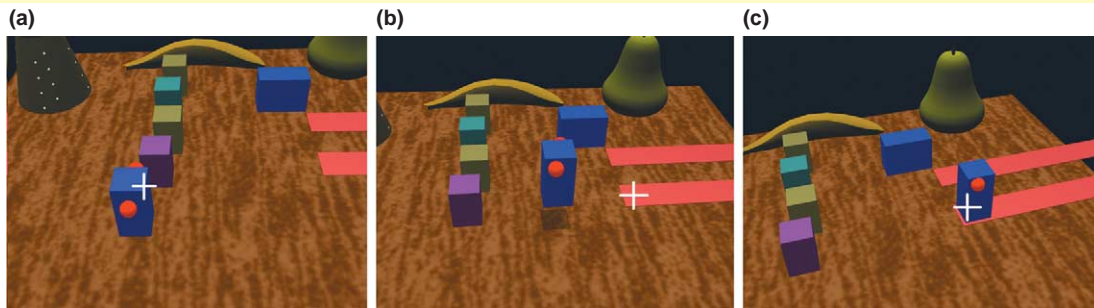

Figure I. Moving bricks task. View of the virtual workspace as a subject **(a)** picks up, **(b)** carries, and **(c)** places a brick on a conveyor belt. The red dots show the fingers, and the white crosses the subject's fixation point. Adapted from [40].

close to intersections. As a consequence of this, subjects were more likely to notice 'Stop' signs located at intersections as opposed to signs in the middle of a block, suggesting in turn that subjects have learnt that traffic signs are more likely around intersections. At a more detailed level, subjects must learn the optimal location for the information they need. For example, when pouring tea, fixation is located at the tip of the teapot spout [10]. Presumably, flow from the spout is best controlled by fixating this location. Another important way in which eye movements reveal the influence of learning is that they are often *pro-active*; that is, saccades are often made to a location in a scene in advance of an expected event. For example, in Land and MacLeod's investigation of cricket, batsmen anticipated the bounce point of the ball, and more skilled batsmen arrived at the bounce point about 100 ms earlier than less skilled players [33]. These saccades were always preceded by a fixation on the ball as it left the bowler's hand, showing that batsmen use current sensory data in combination with learnt models of the ball's motion to predict the location of the bounce. Thus, eye movement patterns appear to be shaped by learnt internal models of the dynamic properties of the world.

Context-dependent neural activity: evidence for visual routines

The eye tracking studies show that the locus and sequence of gaze positions, and the associated visual computations, or visual routines, are orchestrated by the ongoing cognitive goals. The new challenge is to understand how this can be achieved by the neural machinery. There are

two ways in which recent neurophysiological evidence helps us understand the cognitive control of eye movements. The first is the growing evidence that the neurons themselves behave in a task specific manner, even in early cortical visual areas such as V1 (presumably as a consequence of feedback from higher cortical areas). For example, Gilbert and colleagues have shown that the perceptual judgment required of the animal determines the way in which stimulus context modulates the classical receptive field of V1 neurons [43]. Another compelling experiment that reveals this task-dependent processing in primary visual cortex is that by the Roelfsema group, who showed that monkeys solve a visual line tracing task by activating specific cells in striate cortex [38]. These experiments demonstrate that even primary visual cortex can be dynamically reconfigured to selectively extract the specific information required for the momentary task, in a way suggested by the psychophysical data. Higher cortical areas, such as dorso-lateral prefrontal cortex, where cells appear to code the conjunction of specific visual qualities with the learned motor response, also respond in a task specific manner, shaped by experience [44,45]. Thus, the demands of the task seem to be an intrinsic component of the brain's representational structure of visual information.

Eye movements are driven by prospects of reward

The second development in neurophysiology that helps us understand the cognitive control of eye movements is that of showing the influence of reward on eye movements. We know the brain must learn without an explicit programmer, but how is this done? Much research supports a
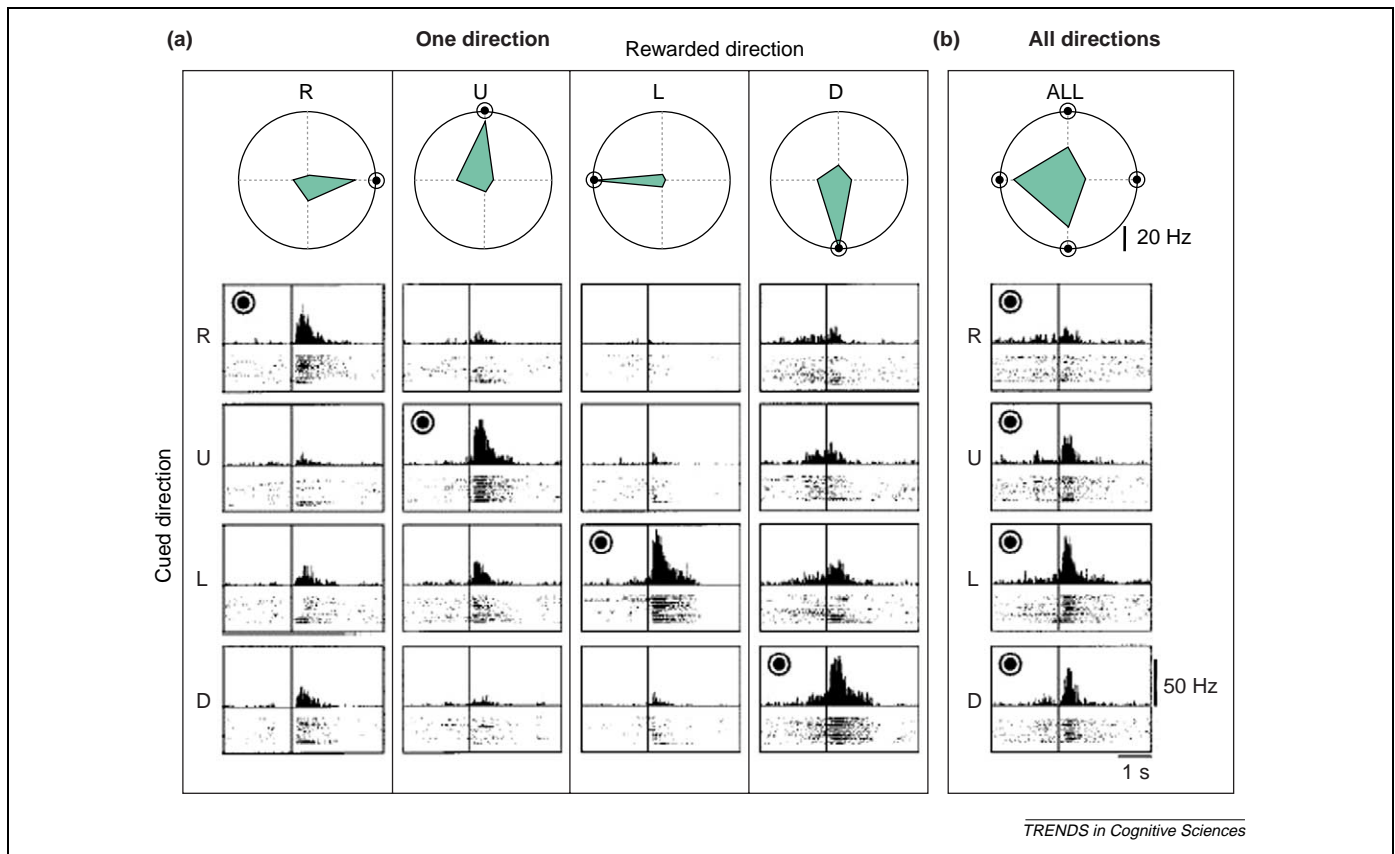
Figure 3. Experiment and model of reward-based eye fixations. Experiments in monkeys by Kawagoe *et al*. [62] show that neurons in the caudate nucleus are very sensitive to reward. In a memory-guided saccade task only one of four directions was rewarded in certain blocks of trials and the rewarded direction is reflected in the tuning functions of neurons responding to that direction. In a control where all directions were rewarded the tuning bias was markedly less directional.

reward-based learning mechanism. The brain has some way of generating a small number of viable alternative behaviors from a vast stored repertoire, and these can be scored with a secondary reward signal believed to be dopamine. The most impressive evidence for this comes from experiments by Schultz that show that dopaminergic neurons in the substantia nigra pars compacta – a part of the basal ganglia system – behave in ways predicted by mathematical models of reinforcement [20,46]. This reward system is integral to the generation of saccadic eye movements. Cortical saccade-related areas (frontal eye fields, dorso-lateral pre-frontal, and lateral intra-parietal) all converge on the caudate nucleus in the basal ganglia, and the cortical-basal ganglia-superior colliculus circuit appears to regulate the control of fixation and the timing of planned movements. This is achieved by regulation of tonic inhibition exerted by the substantia nigra pars reticulata on the superior colliculus, the mid-brain site of saccade generation. Such regulation is a crucial component of task control of fixations. Hikosaka and colleagues have demonstrated that caudate cell responses reflect both the target of an upcoming saccade and the reward expected after making the movement [18,47] (see Figure 3). Thus the neural substrate for learning where to look in the context of a task is present in the basal ganglia.

Reward-based circuitry is widespread
Sensitivity to reward is manifest throughout the saccadic eye movement circuitry. In the lateral intra-parietal area (LIP), the neurons involved in saccadic targeting respond in a graded manner to both the amount of expected reward, and the probability of a reward, in the period before execution of the response [16,17,48,49]. Sensitivity to both these variables is crucial for linking fixation patterns to task demands and is predicted by game theory and decision theory. Cells in the supplementary eye fields also signal the animal's expectation of reward and monitor the outcome of saccades [19]. Sensitivity to stimulus probability is also revealed in 'build-up' neurons in the intermediate layers of the superior colliculus – the mid-brain saccade generator. Basso and Wurtz [50] showed that firing in these neurons before the saccade is proportional to probability that the upcoming target will appear in that location. This forms a nice parallel with psychophysical observations showing that saccade reaction time is similarly influenced by stimulus probability [51]. Hand movements are also sensitive to reward structure. Trommershauser and colleagues [52] have shown that subjects making fast hand movements are able to learn a complicated spatially distributed target reward system and behave in a nearly optimal manner to maximize reward.

Although dopamine is acknowledged to be the main reward signal, attention has turned to the other neuro-transmitters as having a role in second order reward statistics such as uncertainty and risk [53,54]. These statistics are used by Sprague and Ballard [23] to choose between ongoing competing tasks. In their model,

uncertainty increases (together with an attendant cost) when gaze is withheld from an informative scene location, and fixation is allocated to the task with the most to lose. They show that such a cost is calculable within the reinforcement learning framework.

Modeling gaze control in complex tasks with virtual humans

Although the neurophysiological data showing context-specific responses and the importance of reward constitute a crucial substrate for explaining task-directed eye movement patterns, it is still an enormous challenge to understand how these elemental processes are organized to compose the complex gaze patterns observed in everyday behaviors, like those described above. Fortunately there has been a third avenue of development in virtual reality (VR) graphics environments that helps make this link. Such environments now run in real time on standard computing platforms. The value of VR is that it allows the creation of human models that implement complex visuo-motor loops for the control of temporally extended behaviors. Visual input can be captured from the rendered virtual scene, analyzed and the results used to generate motor commands that direct the graphical representation of the virtual agent's body [22]. In this way they help make explicit the set of computational problems involved in generating simple human behaviors. Virtual humans have the advantages of experimental reproducibility, minimal hardware requirements, flexibility, and ease of programming. What these simulations reveal is the different kinds of computational problems that are involved in understanding the brain's allocation of the body's resources. Resource allocation has been extensively studied under the rubric of 'attention' and 'working memory' in psychological experiments. Virtual human simulations provide a complementary approach, wherein a complete system model is produced that includes precise specifications of the ongoing allocation of resources to tasks as revealed through the control of gaze.

One of the principal findings of these simulations is that the information from fixations supports behavior in different ways. In this respect is very helpful to think of the visual computations that need to be done as hierarchically organized, a viewpoint that has not been as obvious from the attention/working-memory dichotomy. Table 1 shows the basic elements of a representative hierarchy that has three levels – Behavior, Arbitration and Context – highlighting the different roles of vision at each level.

(1) *Behavior.* At the level of individual behaviors, the gaze point enables the computation of state information necessary for meeting behavioral goals. (For example, is there an obstacle in the path?) Almost invariably, the visual computation needed in a task context is vastly simpler than that required by general-purpose vision and, as a consequence, can be done very quickly [55,56]. Psychophysical evidence for such context-specific visual computations within a fixation was described above (Box 2), and context-specific neural activity in visual cortex [38,43–45] provides a neural basis for such state computations.

(2) *Arbitration.* At the arbitration or resource allocation level, the principal issue for eye movement control is that the center of gaze (or attentional capacity) is not easily shared and instead must be allocated sequentially to different locations. (For example, different fixation locations are needed for obstacle avoidance and controlling heading direction.) Because the set of active behaviors must share perceptual and physical resources, there must be some mechanism to arbitrate when they make conflicting demands. Task arbitration, particularly dual-task performance, [57,58] has been extensively studied but not from the perspective of gaze allocation. One arbitration model, described below, [23] shows how gaze allocations might be selected to minimize the risk of losing the reward associated with a given behavior.

(3) *Context.* The context level maintains an appropriate set of active behaviors from a much larger library of possible behaviors, given the current goals and environmental conditions. (For example, when reaching an intersection, a walker must detect the new context and choose from a different set of behaviors, such as stopping at the curb.) Thus the function of vision at this level is to achieve an appropriate balance between agenda-driven and environmentally driven visual processing demands. This issue has not received much attention in the cognitive literature but has been extensively studied in robotic models [59–61].

The three hierarchical levels are illustrated in the 'Walter' humanoid simulation conducted by Sprague and Ballard [23] (see Figure 4). In this simulation the direction of gaze is entirely controlled by rewards that are learned en route. As shown in the figure, the issues surrounding eye movements are very different depending on the abstraction level. By making the different level of the hierarchy explicit, the analysis of complex behavior involving eye movements, attention, and working memory, is simplified because the computations at each level are much more independent of each other. Thus different models of any one aspect of the design can be tested in an environment that captures much of the complexity of real human eye movements.

Conclusion

There is a long history of attempts to use eye movements to infer cognitive processes. Saccades are quintessentially voluntary movements [2], and a variety of psychophysical and imaging studies support the idea that the shifts in attention made by the observer are usually reflected in the fixations [6]. Although many fundamental questions remain (see Box 3) the most recent research extends these ideas in several ways. Portable eye tracking technology allows us to study the effects of task requirements on the underlying cognitive process over much more extended periods. Developments in virtual environments allow us to pinpoint the information used moment by moment in a task and reveal it to be exquisitely specific. Neural recordings in monkeys engaged in specific tasks support such computational specificity. The mental programs that drive gaze selection are increasingly seen to be reward-based. This direction is spurred by extensive neural recordings that show vast areas of the brain's gaze computation system exhibit sensitivity to reward signals, as well as developments in reinforcement learning theory that make very specific predictions, some of which

Table 1. Different uses of gaze at different levels of abstraction

Abstraction level	Problem being addressed	Role of vision	Attention	Working memory (WM)
Behavior	A visual routine needs to get state information	Provide state estimation from visual image	Yes	The contents of WM
Arbitration	Different behaviors can have competing demands for the eyes	Move gaze to the best location	Yes	The referents of WM
Context	Current set of behaviors is inadequate	Test for off-agenda exigencies and provide a updated set of behaviors	Yes	–
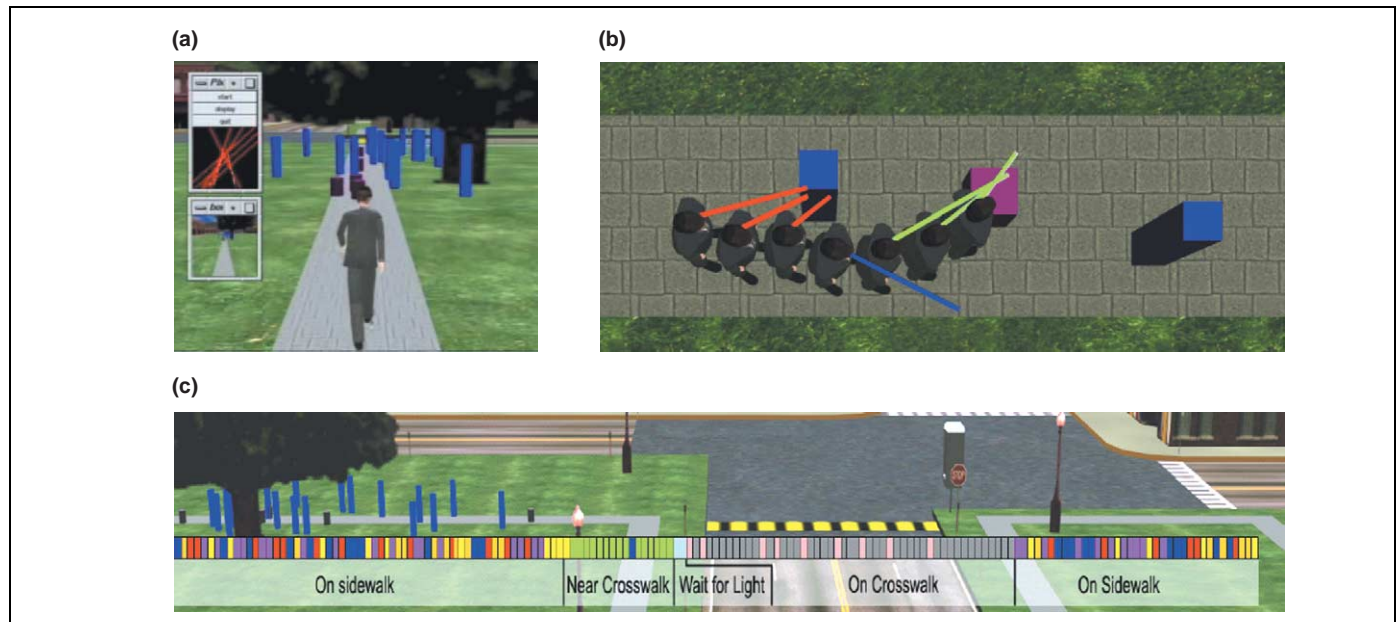
Figure 4. Different gaze functions illustrated in a virtual reality simulation. Three different hierarchical functions of gaze (also shown in Table 1) are shown in a graphic simulation that uses a human-like figure 'Walter' who has learnt to walk along a path in a virtual environment while avoiding obstacles and picking up litter. The model is based on the use of reward-guided visual behaviors that use vision exclusively for perception. **(a) Behavior level:** Walter is shown using a visual routine for determining his location on the sidewalk. The top left hand inset shows the result of a boundary detector looking for sidewalk edges. **(b) Arbitration level:** a delayed exposure shows how Walter allocated gaze in the service of three competing ongoing behaviors (sidewalk following, obstacles, and litter) that need gaze to update internal models. Gaze is given to the behavior with the most to lose as measured by reward expectation computed from reward tables. Initially, the proximity of an obstacle attracts three successive fixations (red gaze vectors), subsequently the sidewalk following behavior is given a fixation (blue) followed by three successive litter pickup fixations (green). **(c) Context level:** depending on the current scene context, Walter activates different sets of behaviors from a larger repertoire. The fixations from the different behaviors are indicated by different colors. For example, when in the 'On Crosswalk' context, the behaviors are twofold: (i) following the crosswalk (gray fixations) and (ii) checking for the other side (pink fixations).

---

have been confirmed. Lastly, reward-based gaze control has been given impetus from graphic simulations. These allow the modeling of situated gaze with vastly increased fidelity and confirm the viability of the understanding of why we look at places in the world in terms of potential rewards.

References
1 Yarbus, A. (1967) *Eye Movements and Vision*, Plenum Press
2 Kowler, E. (1991) The role of visual and cognitive processes in the control of eye movement. In *Eye Movements and Their Role in Visual and Cognitive Processes. Reviews of Oculomotor Research* 4 (Kowler, E. ed.), pp. 1–70, Elsevier
3 Viviani (1991) Eye movements and their role in visual and cognitive processes. *Reviews of Oculomotor Research* (Vol. 4) (Kowler, E. ed.), pp. 1–70, Elsevier
4 Henderson, J. (2003) Human gaze control during real-world scene perception. *Trends Cogn. Sci.* 7, 498–504

5 Liversedge, S. and Findlay, J. (2000) Saccadic eye movements and cognition. *Trends Cogn. Sci.* 4, 6–14

6 Findlay, J. and Gilchrist, I. (2003) *Active Vision*, Oxford University Press

7 Henderson, J. and Ferreira, F. (eds) (2004) *The Interface of Language, Vision, and Action*, Psychology Press

8 Land, M.F. and Lee, D.N. (1994) Where we look when we steer. *Nature* 369, 742–744

9 Land, M. and Furneaux, S. (1997) The knowledge base of the oculomotor system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 352, 1231–1239

10 Land, M.F. *et al*. (1999) Eye movements and the roles of vision in activities of daily living: making a cup of tea. *Perception* 28, 1311–1328

11 Pelz, J.B. and Canosa, R. (2001) Oculomotor behavior and perceptual strategies in complex tasks. *Vision Res.* 41, 3587–3596

12 Hayhoe, M. *et al*. (2003) Visual memory and motor planning in a natural task. *J. Vis.* 3, 49–63

13 Land, M. (2004) Eye movements in daily life. In *The Visual Neurosciences* (Vol 2) (Chalupa, L. and Werner, J., eds), pp. 1357–1368, MIT Press

14 Pelz, J.B. *et al*. (2000) Portable eyetracking: a study of natural eye movements. In *Proceedings of the SPIE, Vol. 3959: Human Vision and Electronic Imaging*, pp. 566–583, SPIE

15 Babcock, J. and Pelz, J. (2004) Building a lightweight eyetracking headgear. *Proceedings ACM SIGCHI: Eye Tracking Research and Applications*, pp. 109–114, ACM Press

16 Platt, M.L. and Glimcher, P.W. (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238

17 Glimcher, P. (2003) The neurobiology of visual-saccadic decision making. *Annu. Rev. Neurosci.* 26, 133–179

18 Hikosaka, O. *et al*. (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 953–978

19 Stuphorn, V. *et al*. (2000) Performance monitoring by the supplementary eye field. *Nature* 408, 857–860

20 Schultz, W. (2000) Multiple reward signals in the brain. *Nat. Rev. Neurosci.* 1, 199–207

21 Sutton, R. and Barto, A. (1998) *Reinforcement Learning*, MIT

22 Terzopoulos, D. and Rabie, T.F. (1997) Animat vision: Active vision in artificial animals. *Videre: J. Comput. Vision Res.* 1, 2–19

23 Sprague, N. and Ballard, D. (2003) Eye movements for reward maximization. In *Advances in Neural Information Processing Systems* (Vol. 16), MIT Press

24 Mannan, S. *et al*. (1997) Fixation patterns made during brief examination of two-dimensional images. *Perception* 26, 1059–1072

25 Itti, L. and Koch, C. (2001) Computational modeling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203

26 Parkhurst, D. *et al*. (2002) Modeling the role of salience in the allocation of overt visual attention. *Vision Res.* 42, 107–123

27 Epelboim, J. *et al*. (1995) The function of visual search and memory in sequential looking tasks. *Vision Res.* 35, 3401–3422

28 Ballard, D. *et al*. (1995) Memory Representations in natural tasks. *Cognitive Neuroscience* 7, 66–80

29 Radach, R. *et al*. Attention, saccade programming, and the timing of eye movement control. *Behav. Brain Sci.* (in press)

30 Deubel, H. and Schneider, W.X. (2000) Saccades, attention, and subjective gaze. *Perception* 29 (Suppl.), 2

31 Turano, K. *et al*. (2003) Oculomotor strategies for the direction of gaze tested with a real-world activity. *Vision Res.* 43, 333–346

32 Land, M. and Hayhoe, M. (2001) In what ways do eye movements contribute to everyday activities? *Vision Res.* 41, 3559–3566

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

34 Rothkopf, C.A. and Pelz, J.B. (2004) Head movement estimation for wearable eye tracker. In *Proceedings ACM SIGCHI: Eye Tracking Research and Applications*, pp. 123–130, ACM Press

35 Johansson, R. *et al*. (2001) Eye-hand coordination in object manipulation. *J. Neurosci.* 21, 6917–6932

36 Hayhoe, M. (2000) Visual routines: a functional account of vision. *Vis. Cogn.* 7, 43–64

37 Ullman, S. (1984) Visual routines. *Cognition* 18, 97–157

38 Roelfsema, P. *et al*. (2000) The implementation of visual routines. *Vision Res.* 40, 1385–1411

39 Hayhoe, M. *et al*. (1998) Task constraints in visual working memory. *Vision Res.* 38, 125–137

40 Triesch, J. *et al*. (2003) What you see is what you need. *J. Vis.* 3, 86–94

41 Chapman, P. and Underwood, G. (1998) Visual search of dynamic scenes: event types and the role of experience in viewing driving situations. In *Eye Guidance in Reading and Scene Perception* (Underwood, G., ed.), pp. 369–394, Elsevier

42 Shinoda, H. *et al*. (2001) Attention in natural environments. *Vision Res.* 41, 3535–3546

43 Gilbert, C. *et al*. (2001) The neural basis of perceptual learning. *Neuron* 31, 681–697

44 Assad, W. *et al*. (1998) Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21, 1399–1407

45 Miller, E. and Cohen, J. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202

46 Montague, P.R. *et al*. (2004) Computational roles for dopamine in behavioral control. *Nature* 431, 760–767

47 Watanabe, K. *et al*. (2003) Neural correlates of rewarded and unrewarded movements in the primate caudate nucleus. *J. Neurosci.* 23, 10052–10057

48 Dorris, M.C. and Glimcher, P.W. (2004) Activity in posterior parietal cortex is correlated with the subjective desireability of an action. *Neuron* 44, 365–378

49 Sugrue, L.P. *et al*. (2004) Matching behavior and the encoding of value in parietal cortex. *Science* 304, 1782–1787

50 Basso, M. and Wurtz, R. (1998) Modulation of neuronal activity in superior colliculus by changes in target probability. *J. Neurosci.* 18, 7519–7534

51 He, P. and Kowler, E. (1989) The role of location probability in the programming of saccades: implications for center-of-gravity tendencies. *Vision Res.* 29, 1165–1181

52 Trommershauser, J. *et al*. (2003) Statistical decision theory and the selection of rapid, goal-directed movements. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 20, 1419–1433

53 Doya, K. (2002) Metalearning and neuromodulation. *Neural Netw.* 15, 495–506

54 Seymour, B. *et al*. (2004) Temporal difference models describe higher order learning in humans. *Nature* 429, 664–667

55 Ballard, D. *et al*. (1997) Deictic Codes for the Embodiment of Cognition. *Behav. Brain Sci.* 20, 723–767

56 Salgian, G. & Ballard, D.H. (1998a) Visual routines for autonomous driving. *Proceedings of the 6th Int. Conf. Comput. Vision*, pp. 876–882, IEEE computer Society Press

57 Ruthruff, E. *et al*. (2003) Dual-task interference with equal task emphasis: graded capacity-sharing or central postponement? *Percept. Psychophys.* 65, 801–816

58 Lee, D.K. *et al*. (1999) Attentional capacity is undifferentiated: concurrent discrimination of form, color, and motion. *Percept. Psychophys.* 61, 1241–1255

59 Bonasso, R.P. *et al*. (1997) Experiences with an architecture for intelligent, reactive agents. *J. Exp. Theor. Artif. Intell.* 9, 237–256

60 Hartley, R. and Pipitone, F. (1991) Experiments with the subsumption architecture. *Proceedings IEEE Int. Conf. Robotics and Automation* (Vol. 2), pp 1652–1658, IEEE

61 Humphrys, M. (1996) Action selection methods using reinforcement learning. In *From Animals to Animats 4: Proceedings Fourth Int. Conf. on Simulation of Adaptive Behavior* (Maes, P. *et al*., eds), pp. 135–144, MIT Press

62 Kawagoe, R. *et al*. (1998) Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416