

# Adaptive Gaze Control in Natural Environments

Jelena Jovancevic-Misic<sup>1</sup> and Mary Hayhoe<sup>2</sup>

<sup>1</sup>Department of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, Rochester, New York 14627-0270, and <sup>2</sup>Center for Perceptual Systems, University of Texas at Austin, Austin, Texas 78712

The sequential acquisition of visual information from scenes is a fundamental component of natural visually guided behavior. However, little is known about the control mechanisms responsible for the eye movement sequences that are executed in the service of such behavior. Theoretical attempts to explain gaze patterns have almost exclusively concerned two-dimensional displays that do not accurately reflect the demands of natural behavior in dynamic environments or the importance of the observer's behavioral goals. A difficult problem for all models of gaze control, intrinsic to selective perceptual systems, is how to detect important but unexpected stimuli without consuming excessive computational resources. We show, in a real walking environment, that human gaze patterns are remarkably sensitive to the probabilistic structure of the environment, suggesting that observers handle the uncertainty of the natural world by proactively allocating gaze on the basis of learned statistical structure. This is consistent with the role of reward in the oculomotor neural circuitry and supports a reinforcement learning approach to understanding gaze control in natural environments.

## Introduction

The control of the sequential acquisition of information from the environment via gaze changes is a critical problem for any selective perceptual system. A central open question is this: how does the system balance the need to attend to existing goals versus maintaining sensitivity to new information that may pose opportunities or threats? It is often assumed that the solution to the problem of gaze allocation is that attention is attracted by salient stimuli or events in the visual array, such as color, intensity, contrast, and edge orientation (Itti and Koch, 2000; Henderson, 2003). In addition, sudden onset or motion stimuli typically capture attention, indicated by interference with the ongoing task (Theeuwes and Godijn, 2001; Franconeri and Simons, 2003; Lin et al., 2008). However, it seems unlikely that these mechanisms can solve the general problem. Saliency models are inflexible and can account for only a small fraction of the observed fixations in natural behavior (Land, 2004; Hayhoe and Ballard, 2005; Jovancevic et al., 2006; Rothkopf et al., 2007). Most importantly, it is unlikely that the experimental contexts that have been examined reflect the demands of natural visually guided behavior. In natural behavior, observers control both what information is selected from the image and the time at which it is selected. This control mechanism is hard to investigate in standard experimental paradigms in which the experimenter controls these processes. In addition, both the stimulus conditions and task context are quite different. Almost all of the work showing stimulus-based effects of attentional or oculomotor capture has been done with two-dimensional experimental displays and either simple geo-

metric stimuli or photographic renderings of natural scenes. Even when images of natural scenes are used, the situation does not accurately reflect the stimulus conditions in visually guided behavior in real, three-dimensional environments, in which a time-varying image sequence is generated on the retina as the observer moves through the scene. Moreover, many kinds of information are important to the subject and need to be attended, not just sudden onsets or other salient regions (for example, irregularities in the pavement or the location of obstacles). Consequently, a mechanism that relies on stimulus properties for attracting attention or gaze is not likely to be reliably effective.

Recent work in natural tasks has demonstrated that the observer's cognitive goals play a critical role in the distribution of gaze during a wide range of natural behaviors (Land and Farnieux, 1997; Land et al., 1999; Hayhoe et al., 2003; Land, 2004; Hayhoe and Ballard, 2005). However, this does not solve the general problem of gaze allocation. When making tea or sandwiches, items remain in stable locations with stable properties, and a subject's behavioral goals can be achieved by fixating task-relevant objects, according to some learned protocol. In other environments, such as driving or walking down the street, the goals are less well defined, and it is not always possible to anticipate what information is required. How do observers distribute gaze appropriately when the information is not entirely predictable? Attentional limitations mean that such stimuli will sometimes be missed, and the problem for the brain is to minimize this probability. The present work attempts to explain the mechanism by which this happens.

One possible solution is that, for a specific context, observers learn the dynamic properties of the world and use these learned models as a basis for distributing gaze and attention where they are likely to be needed. Although the importance of learning in controlling gaze is implicit in the task dependence of fixation patterns (Land, 2004; Hayhoe and Ballard, 2005), just how such learning might function to control gaze in dynamic, unpredict-

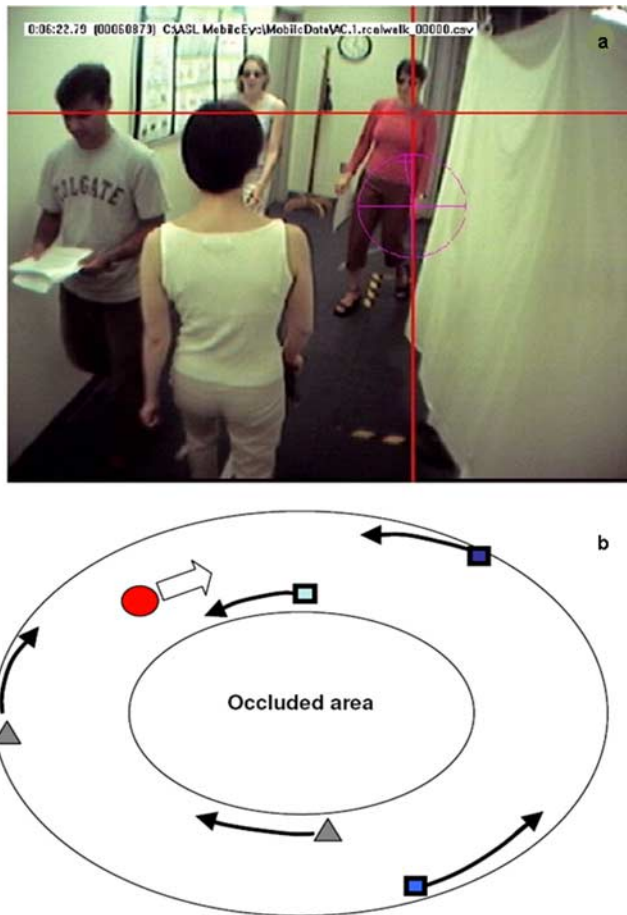
Received Nov. 20, 2008; revised March 25, 2009; accepted April 2, 2009.

This work was supported by National Institutes of Health Grants EY05729 and RR09283. We thank K. Chajka for assistance with the experiments.

Correspondence should be addressed to Mary Hayhoe, Center for Perceptual Systems, 1 University Station, #A8000, University of Texas at Austin, Austin, TX 78712. E-mail: mary@mail.cps.utexas.edu.

DOI:10.1523/JNEUROSCI.5570-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/296234-05\$15.00/0



**Figure 1.** *a*, Subject's view of the path from a camera mounted on the head. The crosshair shows gaze. *b*, Illustration of a path around the occluded area: the red circle represents the subject, the blue squares represent the pedestrians walking in the direction opposite to that of a subject, and the gray triangles represent pedestrians walking in the same direction as the subject.

able environments, and the properties of that learning, are essentially unexplored.

To investigate whether human gaze patterns reflect sensitivity to the statistical properties of the environment, we attempted to actively manipulate gaze allocation by varying the probability of potential collisions. The hypothesis explored in the current study is that attentional or oculomotor capture is unnecessary if subjects learn the statistical properties of the environment, such as how often pedestrians need to be fixated to monitor their direction.

## Materials and Methods

We devised an experiment in which subjects walked along a path in the presence of other pedestrians who behaved in characteristic ways. Different pedestrians had different probabilities of heading toward the subject for a brief period. Detecting such potential collisions is a situation in which attentional capture mechanisms might be useful. However, we have shown in previous work (Jovancevic et al., 2006) that such automatic capture is probably not effective in this context, despite its effectiveness in simpler displays (Lin et al., 2008).

Observers walked along an oval path in a large room, around a central structure that restricted view of the other side of the room. Gaze was tracked with an ASL infrared, video-based, portable eye tracker mounted on the head. Figure 1 shows the subject's view from a head-mounted camera, with eye position indicated by the cross. System accuracy was  $\sim 0.5^\circ$ . Some of the oncoming walkers were instructed to veer toward the

subject for an  $\sim 1$  s period, at a distance of  $\sim 4$  m from the subject. Walkers always returned to their original path before an actual collision occurred. To investigate whether subjects learn how often to fixate particular regions in the natural world (in this case, other pedestrians), three of the walkers followed a script. One walker (Rogue) went on a collision path every time he/she approached the subject, a second walker went on a collision path on half the occasions (Risky), and a third never collided (Safe). Although no actual collisions ever occurred, for simplicity, we refer to the event of going on a collision course as a "collision" and the pedestrians as "colliders." The subjects were not informed about these manipulations and were simply told to walk around the path for a number of laps and to avoid the other pedestrians.

Subjects were instructed to walk at a normal pace around the path for four sets of 12 laps, separated by a short break and calibration check. In addition to the subject, there were five pedestrians: three walking in the direction opposite and two walking in the same direction as the subject. In another condition, instead of veering toward the subject, Rogue and Risky pedestrians stopped for 1 s and then resumed walking. After each trial of 12 laps, the Rogue and Safe pedestrians changed roles, whereas the Risky pedestrian stayed the same. Another condition in which there were no unexpected events (potential collisions or stopping) was introduced as a control. Different starting points of the pedestrians and their different speeds ensured that they were evenly distributed around the central occluder and that their layout on the path was constantly changing. Thus, the subject never saw the same layout twice.

Analysis of fixations on pedestrians was performed on all five subjects in each condition (a total of 30 subjects). Fixations were determined using manual frame-by-frame coding from the 30 Hz video record. To allow for the noise in the eye tracking signal, fixations were scored as being on the pedestrian if they fell within a  $\sim 0.75^\circ$  window on either side of the pedestrian. Because the pedestrians typically displaced laterally with respect to the subject during the period gaze was maintained, we treated a "fixation" as the period when gaze remained at the same location with respect to the pedestrian, although it included some smooth rotational component and is more properly specified as "gaze." For every encounter, a record was made of fixations on pedestrians after they first appear in the field of view. Fixation probability was calculated from the number of laps in which a particular pedestrian got a fixation (at any point during the lap) relative to the total number of laps (additional fixations during the same lap were not counted). The period a pedestrian was visible was variable and differed between veering and stopping conditions by  $\sim 200$  ms. We did not normalize for this. However, because most are fixations concentrated in the period soon after the pedestrian becomes visible, this is unlikely to affect the data significantly. For Rogue and Risky pedestrians, this also constituted the majority of colliders fixated, because most of the initial fixations on pedestrians continued into the collision period. To investigate the time course of learning, we analyzed gaze durations, that is, the total time spent fixating a pedestrian per encounter as well as latencies of a first fixation on a pedestrian after its appearance in the field of view as a function of time across a trial. Mean latencies were computed over laps in which a fixation occurred.

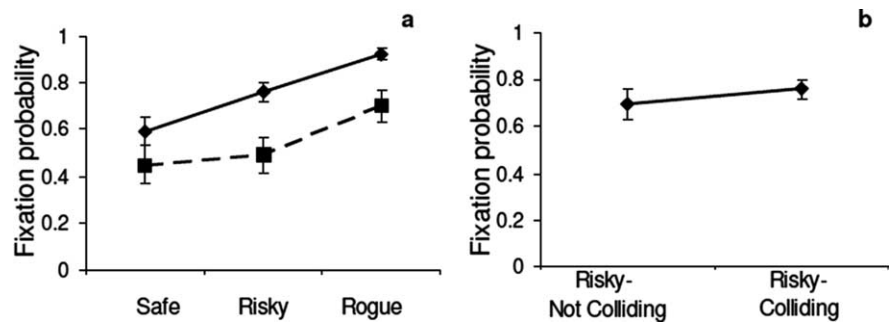
## Results

Subjects clearly increase fixation probability depending on (potential) collision probability, as shown in Figure 2*a*. Subjects fixated the Safe pedestrian least and the Rogue pedestrian most. These differences were significant (one-way ANOVA,  $F_{(2,12)} = 14.426$ ;  $p = 0.00064$ ) as were *post hoc* comparisons on all pairs of means, both within the colliding condition, as well as comparisons of fixated pedestrians between colliding and stopping conditions. This would be less interesting if the fixations were simply evoked by the visual stimulus accompanying the path deviation. However, fixation probabilities were calculated for the entire period the pedestrian was present in the subject's field of view, typically  $\sim 2$  s, not just during the 1 s collision period. Figure 2*b* shows that the fixation probability on the Risky walkers is almost identical whether or not a collision event actually occurs on that

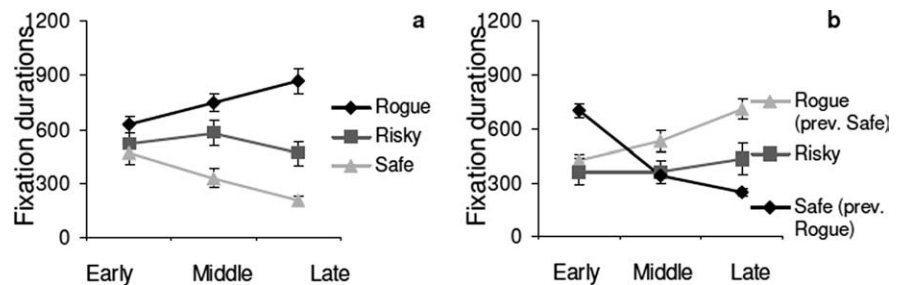
lap. [The small difference was not significant ( $p = 0.42$ ).] Thus, the probability of fixation depends more on the pedestrian's characteristic behavior than the collision event itself on any given trial. Most of the fixations began before the pedestrian started veering. In the case of the Rogue, 92% of fixations occurred before the onset of the collision course. Thus, the burden of detection is carried by proactive deployment of gaze, presumably as a result of recognition of the particular pedestrian using the peripheral retina. Fixating a pedestrian allows the subject to monitor the walker's direction, so that detection of a path change results from the fact that the walker is already being fixated rather than attentional capture by an untoward event such as a path deviation. Although we have no separate measure of speed or path deviations, subjects did not appear to modify their path in response to the veering pedestrians (Jovancevic et al., 2006). Indeed, because the veering always occurred at  $\sim 4$  m distance from the subject, such a modification was not necessary.

To see how quickly a subjects' gaze patterns adjust to pedestrian behavior, we divided the 12 laps into early, middle, and late periods, each consisting of four laps, and four or five encounters with each pedestrian. Because it is difficult to measure fixation probability reliably with so few encounters, we measured total time the pedestrian was fixated per encounter. (If more than one fixation occurred, they were summed.) We also attempted to manipulate gaze by changing the pedestrians' characteristic behavior. After subjects had completed 12 laps around the path (trial 1), they were then required to complete a second set of 12 laps (trial 2) in which the Rogue and Safe pedestrians switched roles. The Risky pedestrian stayed the same. Gaze duration on the different pedestrians as a function of time is shown in Figure 3. The left side of the figure shows data from the first 12 laps, and the right side shows data from the second 12 laps, after the pedestrians changed behavior. Initially, all pedestrians are fixated for  $\sim 500$  ms. After only approximately four encounters, gaze duration on Rogue and Safe pedestrians have diverged substantially, and, by the last four laps, subjects spend  $\sim 900$  ms fixating the Rogue and only  $\sim 200$  ms fixating the Safe pedestrians. This change in gaze durations was statistically significant for Rogue ( $F_{(1,12)} = 4.61$ ;  $p = 0.033$ ) and Safe ( $F_{(2,12)} = 7.61$ ;  $p = 0.007$ ) pedestrians. Time spent fixating the Risky pedestrian remains approximately constant. After the Rogue and Safe pedestrians changed behaviors, fixation time changes according to the pedestrians' new behavior, although it does not completely return to the original value over the course of the 12 laps.

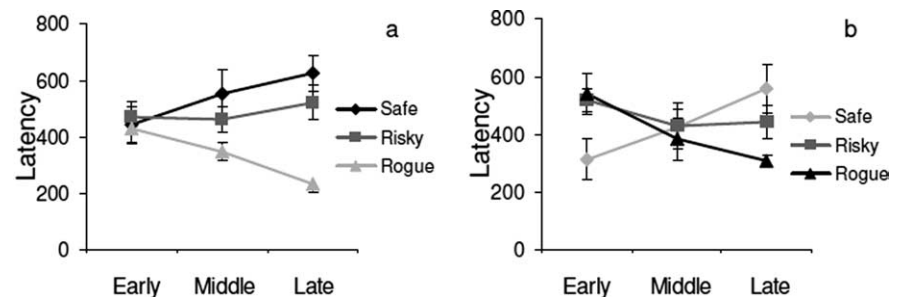
A similar pattern occurs for latency, shown in Figure 4. When pedestrians come into the field of view, they are initially fixated with a latency of  $\sim 450$  ms. By the end of the trial, Rogues are fixated within 200 ms, whereas the fixation latency for the Safe



**Figure 2.** *a*, Effect of collision probability on fixations on pedestrians. Fixation probability was calculated over the entire period the pedestrian was visible during each lap. Filled line, Colliding condition; dashed line, stopping condition (see Results). *b*, Effect of collision event on fixations. Probability of fixating the Risky pedestrian on laps when it was involved in a collision event and when it was not. Error bars in all plots are  $\pm 1$  SEM across subjects, five subjects per data point.



**Figure 3.** Time course of learning. Change in (summed) gaze durations within 12 lap sets divided into three, four-lap periods labeled Early, Middle, Late. *a*, First 12 laps, no previous experience. *b*, Second set of 12 laps, after conflicting experience.



**Figure 4.** Time course of learning. Gaze latency after pedestrian appearance in the field of view. *a*, First 12 laps, no previous experience. *b*, Second set of 12 laps, after conflicting experience. Error bars are  $\pm 1$  SEM across subjects;  $n = 5$ .

pedestrian increased to  $\sim 600$  ms. These changes were statistically significant for Rogue ( $F_{(2,12)} = 6.59$ ;  $p = 0.012$ ) but not for Safe ( $p = 0.24$ ). Subjects appear to use gaze in a proactive manner, by fixating the Rogue early and maintaining gaze for an extended period. Thus, subjects quickly learn to adjust gaze priorities to environmental probabilities. Again, latencies do not quite adjust to the previous values after the change in pedestrian behavior within the 12 laps, but the rate of change is comparable. Gaze control appears to be remarkably sensitive to changes in the current environmental statistics, and this presumably accounts for much of its effectiveness in attending to the right thing at the right time.

It is important to note that, in all the figures, the error bars reflect SEMs between subjects and are calculated across only five subjects for any particular point. Given that there were no explicit instructions, it is remarkable that subjects behave in such a similar manner, as indicated by the small SEs. This suggests that the behavior we observe, and the rapid adjustment of gaze probabilities, reflect a stable and lawful property of natural gaze behavior.



We have so far assumed that the regions people look at in the scene reflect the needs of the particular task context, as indicated by previous observations of natural tasks (Itti and Koch, 2000; Henderson, 2003). The fact that gaze allocation adjusts rapidly with the probability of potential collisions supports this assumption. However, to strengthen the claim that the observer's behavioral goals dictate gaze behavior rather than the properties of the stimulus per se, we performed a similar experiment with different subjects, in which pedestrians merely stopped rather than going on a collision course. All other aspects of the experiment were the same. The visual stimulus associated with stopping should be comparably salient to that associated with a potential collision. Technically, saliency is only defined with respect to some model, such as that of Itti and Koch (2000). Because either a collision course or a pedestrian stopping involves a change in the pattern of motion with respect to the surroundings, in a large retinal region, we are assuming that such events would be salient according to such models. (A collision course involves radial expansion of the pedestrian. Stopping involves translation and expansion, although at a slower rate.) There is substantial evidence that the looming stimulus caused by a collision course is highly salient. This issue was discussed further by Jovancevic et al. (2006). However, the collision is more potentially dangerous to the walker than stopping. A comparison of the two events is shown in Figure 2*a*. The probability of fixating pedestrians is shown for the first set of 12 laps. Both probabilities are lower when the pedestrian simply stops. On average, pedestrians are fixated ~15% less when stopping as opposed to colliding. This reduction was found to be statistically significant ( $F_{(1,8)} = 8.61$ ;  $p = 0.019$ ). Presumably, it is relevant for the observer to know what the other pedestrians are doing, so they are accorded some gaze time in the stopping condition, although less than with path deviations. Note that the Safe pedestrian never stops or collides but is still fixated less in the Stopping condition, in which only the behavior of the Rogue is different. [The difference between means was not significant in trial 1, but was significant in Trial 2 ( $F_{(1,8)} = 7.43$ ;  $p = 0.026$ ).] This suggests that the behavior of the Rogue pedestrian influences fixations on all pedestrians. It also suggests that the reduction in fixations is caused by the change in behavioral relevance, not by reduced salience of the stopping event relative to collisions, because the Safe pedestrian does not in fact change his behavior.

## Discussion

In summary, the experiment reveals a number of important determinants of gaze in dynamic environments such as walking. We showed that manipulation of the probability of a potential collision by a risky pedestrian is accompanied by a rapid change in gaze allocation. Subjects learn new priorities for gaze allocation within a few encounters and look both sooner and longer at potentially dangerous pedestrians. It has been demonstrated that visual search in two-dimensional images of natural scenes is influenced by learned scene statistics (Eckstein et al., 2006; Neider and Zelinsky, 2006; Torralba et al., 2006). The present experiment reveals that observers are also sensitive to the dynamic structure of natural environments and rapidly adjust their behavior depending on the statistical properties of the current environment. In addition, we have demonstrated that the learned properties of the environment are used to predict future state and allocate gaze proactively. We also observed that fixation probabilities are modulated by behavioral significance of the pedestrians' behavior, with colliding pedestrians being fixated more often than stopping pedestrians. Gaze allocation also depends on the

set of tasks the observer must attend to, so that pedestrians must have some strategy for allocating gaze between different tasks. Although it might be argued that it is not surprising that observers learn to fixate the potentially colliding pedestrians, such rapid adaptation of gaze selection has not been demonstrated previously. What is of interest here is the quantitative aspects of performance. Not only were fixation probabilities very similar between subjects, but so, too, were learning rates, fixation durations, and fixation latencies, despite the complexity of walking in a real, dynamic environment and the lack of explicit instructions. This suggests that learned environmental statistics, behavioral significance or risk, and competing tasks are important determinants of gaze in many natural contexts.

How can we explain these effects? Currently, no model of gaze control can account for control of gaze in natural, dynamic environments. As discussed above, the class of models based on stimulus properties (i.e., saliency models) cannot explain the gaze patterns we observe. Although there have been several attempts to model top-down effects in gaze control, such attempts usually take the form of a top-down weighting of the saliency map provided by bottom-up filtering of the image (Wolfe, 1994; Navalpakkam and Itti, 2006; Torralba et al., 2006) and do not address the mechanisms that determine whether, and when, an object or location is chosen by the observer to be a target in the context of an ongoing behavioral sequence. Senders (1983) proposed an information theoretic model for distributing gaze between time-varying instruments on a display panel, a problem that shares some of the elements of the current situation. However, the model of Sprague et al. (2007) comes closest to being able to handle the observed behavior. In their model, a simulated walker has three tasks: to follow the sidewalk, to avoid obstacles, and to pick up objects. The appropriate schedule for allocation of attention or gaze to each task must be learned by experience with the frequency of obstacles, how often it is necessary to look at the path, etc. To choose between ongoing competing tasks such as avoiding obstacles and controlling direction of locomotion, in their model, uncertainty increases (together with cost) when gaze is withheld from an informative scene location. Fixation is allocated to the task that would have the greatest cost if the relevant information were not updated. They show that such a cost is calculable within a reinforcement learning framework. The development of such models provides a formal system for modeling complex behavioral sequences, a critical component in being able to go beyond a simple description of fixation patterns as "task driven," to being able to predict the observed fixation sequences in natural behavior (Rothkopf et al., 2007). The environment for the model of Sprague et al. (2007) is simpler than many natural scenes in that it does not involve dynamic obstacles. However, it does show how gaze behavior can be learned and embodies the principles of fixation patterns based on risk and competition between different task components.

Another critical development in understanding the control of gaze patterns is the finding that many of the regions involved in saccade target selection and generation are sensitive to expectation of reward. Saccade-related areas in the cortex (lateral intraparietal area, frontal eye field, supplementary eye field, and dorsolateral prefrontal cortex) all exhibit sensitivity to reward (Platt and Glimcher, 1999; Stuphorn et al., 2000; Dorris and Glimcher, 2004; Sugrue et al., 2004). These areas 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, by regulating tonic inhibition on the superior colliculus. Such regulation seems

likely to be a critical component of task control of fixations. Caudate cell responses reflect both the target of an upcoming saccade and the reward expected after making the movement (Watanabe et al., 2003). Thus, the neural substrate for learning fixation strategies in the context of a task is potentially available in the basal ganglia (Hikosaka et al., 2006). Supplementary eye fields and anterior cingulate cortex may also play an important behavioral monitoring role (Schall et al., 2002; Stuphorn and Schall, 2006). Although it has not been demonstrated directly that the acquisition of information per se is rewarding, all complex behaviors involve secondary reward of some kind, and the acquisition of information is a critical step in achieving behavioral goals. The present experiments provide evidence for reward sensitivity of the eye movements that occur in natural behavior.

In conclusion, subjects use learned scene statistics of natural environments to guide eye movements proactively, in anticipation of events that are likely to occur in the scene. Because actions such as avoiding a pedestrian take of the order of a second to accomplish, this strategy allows observers to plan their path rather than depending on reactive eye and body movements, which may come too late. The present experiments suggest that prediction based on environmental models is a central component of gaze control and support a reinforcement learning approach to understanding gaze control in natural environments.

## References

- Dorris MC, Glimcher PW (2004) Activity in posterior parietal cortex is correlated with the subjective desirability of an action. *Neuron* 44:365–378.
- Eckstein MP, Drescher BA, Shimozaki SS (2006) Attentional cues in real scenes, saccadic targeting and Bayesian priors. *Psychol Sci* 17:973–980.
- Franconeri SL, Simons DJ (2003) Moving and looming stimuli capture attention. *Percept Psychophys* 65:999–1010.
- Hayhoe M, Ballard D (2005) Eye movements in natural behavior. *Trends Cogn Sci* 9:188–194.
- Hayhoe MM, Shrivastava A, Mruczek R, Pelz JB (2003) Visual memory and motor planning in a natural task. *J Vis* 3:49–63.
- Henderson JM (2003) Human gaze control during real-world scene perception. *Trends Cogn Sci* 7:498–504.
- Hikosaka O, Nakamura K, Nakahara H (2006) Basal ganglia orient eyes to reward. *J Neurophysiol* 95:567–584.
- Itti L, Koch C (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res* 40:1489–1506.
- Jovancevic J, Sullivan B, Hayhoe M (2006) Control of attention and gaze in complex environments. *J Vis* 6:1431–1450.
- Land M (2004) Eye movements in daily life. In: *The visual neurosciences*, Vol 2 (Chalupa L, Werner J, eds), pp 1357–1368. Cambridge, MA: MIT.
- Land M, 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 MF, Furneaux S (1997) The knowledge base of the oculomotor system. *Philos Trans R Soc Lond B Biol Sci* 352:1231–1239.
- Lin JY, Franconeri S, Enns JT (2008) Objects on a collision path with the observer demand attention. *Psychol Sci* 19:686–692.
- Navalpakkam V, Itti L (2006) Top-down attention selection is fine grained. *J Vis* 6:1180–1193.
- Neider MB, Zelinsky GJ (2006) Scene context guides eye movements during visual search. *Vision Res* 46:614–621.
- Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400:233–238.
- Rothkopf C, Ballard D, Hayhoe M (2007) Task and scene context determine where you look. *J Vis* 7:16.1–6.20.
- Schall JD, Stuphorn V, Brown JW (2002) Monitoring and control of gaze by the frontal lobes. *Neuron* 36:309–322.
- Senders JW (1983) Visual scanning processes. PhD thesis, Catholic University Tilberg.
- Sprague N, Ballard D, Robinson A (2007) Modeling embodied visual behaviors. *ACM Trans Action Percept* 4:2.
- Stuphorn V, Schall JD (2006) Executive control of countermanding saccades by the supplementary eye field. *Nat Neurosci* 9:925–931.
- Stuphorn V, Bauswein E, Hoffmann KP (2000) Neurons in the primate superior colliculus coding for arm movements in gaze-related coordinates. *J Neurophysiol* 83:1283–1299.
- Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304:1782–1787.
- Theeuwes J, Godijn R (2001) Attentional and oculomotor capture. In: *Attraction, distraction, and action: multiple perspectives on attentional capture* (Folk C, Gibson B, eds), pp 121–150. Amsterdam: Elsevier.
- Torralba A, Oliva A, Castelano MS, Henderson JM (2006) Contextual guidance of eye movements and attention in real world scenes: the role of global features in object search. *Psychol Rev* 113:766–786.
- Watanabe K, Lauwereyns J, Hikosaka O (2003) Neural correlates of rewarded and unrewarded movements in the primate caudate nucleus. *J Neurosci* 23:10052–10057.
- Wolfe JM (1994) Guided search 2.0. A revised model of visual search. *Psychon Bull Rev* 1:202–238.