

Initial saccades predict manual recognition choices in the monkey

David L. Sheinberg^{a,*}, Jessie J. Peissig^a, Keisuke Kawasaki^b, Ryan E.B. Mruzek^a

^a *Department of Neuroscience, Brown University, Providence, RI 02912, USA*

^b *Brain Science Program, Brown University, Providence, RI 02912, USA*

Received 9 January 2006; received in revised form 20 March 2006

Abstract

In animals with specialized foveae, eye position has a direct influence over the acquisition of detailed visual information. At the same time, eye movements executed during natural behaviors are closely linked with motor actions. In this study, we investigated patterns of eye movements during a simple visual discrimination task. Three rhesus monkeys learned to recognize images of real world objects with no explicit constraints on eye position. Analysis of the monkeys' eye movements showed that although the endpoint of the initial saccade depended on the particular visual stimulus, the trajectory of the first saccades also reliably predicted the manual response associated with that stimulus. We thus observed that initial saccades executed in a recognition task reflect both perceptual and motor aspects of a visual task. This pattern of eye movements emerged spontaneously in all three animals tested despite the fact that saccades were never explicitly rewarded. As the average saccade latency was under 200 ms, object specific learned associations must have exerted their influence over the initial saccade even earlier, providing a novel temporal marker for the rapidity of visual recognition processes. Taken together, these results suggest that caution should be exercised when interpreting the meaning of oculomotor patterns observed during perceptual tasks, as these blur the line between perceptual processing and motor preparation.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Saccades; Object recognition; Macaque; Eye-hand coordination; Oculomotor

1. Introduction

Natural vision involves eye movements and visual perception in the absence of eye movements is not a normal occurrence. Eye movements play a direct role in the sensory acquisition of visual information, and links between eye movements and perception (Buswell, 1935; Yarbus, 1967) and cognition (Kinsbourne, 1972; Kinsbourne, 1974) are well established. At the same time, the behavioral output of the oculomotor system—one of the best understood motor control systems (Carpenter, 1988; Robinson, 1981)—has been clearly linked to action planning (Cutting, Alliprandini, & Wang, 2000; Hayhoe, Shrivastava, Mruzek, & Pelz, 2003; Land, Mennie, & Rusted, 1999; Land & McLeod, 2000; Land & Tatler, 2001; Negggers & Bekkering, 2002).

Given this dual role of eye movements in both action and perception, we chose to explore the relationship between saccadic eye movements and recognition in rhesus monkeys. Both human and non-human primates rapidly redirect their gaze to distinct spatial locations in the course of natural visual activities. Observations of these overt gaze shifts have lead many to hypothesize that this rich behavior can provide evidence about how observers gather information critical for recognition (Buswell, 1935; Noton & Stark, 1971; Tinker, 1946; Yarbus, 1967) and how attention is distributed during active visual processing (Deubel & Schneider, 1996; Kowler, Anderson, Doshier, & Blaser, 1995; Remington, 1980). The level at which saccades are planned remains a matter of considerable debate (Findlay & Gilchrist, 2003; Henderson, 2003). On one hand, saccades are certainly constrained by the visual array, where an image-based saliency map may be critical for attracting the eyes to particular stimulus locations. At the same time, internally

* Corresponding author. Fax: +1 401 863 1074.

E-mail address: David_Sheinberg@brown.edu (D.L. Sheinberg).

guided plans related to task demands (Henderson, Weeks, & Hollingworth, 1999; Yarbus, 1967) or contextual knowledge (Henderson et al., 1999; Oliva, Torralba, Castelhan, & Henderson, 2003) may also directly influence saccade targets. Furthermore, planned actions, such as arm movements, may influence saccade trajectories due to interactions between hand-centered and eye-centered frames of reference (Tipper, Howard, & Paul, 2001).

In the current experiment, we designed a simple recognition task requiring manual responses, during which oculomotor behavior was monitored, but was neither directly rewarded nor artificially constrained. These conditions differ from most previous recognition studies in monkeys in that we did not use eye movements as an operant response and we did not require constrained fixation. Previous experiments suggest that in tasks requiring manual responses, saccades almost always occur in advance of the overt manual responses (Prablanc, Echallier, Komilis, & Jeannerod, 1979). We were interested in how naturally occurring saccades may be influenced by developing motor plans during a visual classification task, and when these saccades occurred relative to stimulus appearance and manual response time. In all three animals tested, we observed that initial saccades executed in our recognition task reflect both perceptual and motor aspects of the visual task. Furthermore, our data indicate that recognition occurred prior to 200 ms after stimulus onset supporting recent models of rapid visual processing (Kirchner & Thorpe, 2006; Thorpe, Delorme, & Van Rullen, 2001).

2. Methods

2.1. Subjects

Three male rhesus monkeys (*Macaca mulatta*), ages 5–9 years and weighing between 9 and 13 kg, were the subjects in this study. Prior to the experiment, the monkeys had been familiarized with the behavioral apparatus and had participated in unrelated studies. The monkeys had a titanium head restraint post-surgically implanted and one of the three had a scleral eye coil (Judge, Richmond, & Chu, 1980; Robinson, 1963). All surgeries were performed using sterile technique while the animals were intubated and anesthetized using isoflurane gas. Procedures conformed to the NRC *Guide for the Care and Use of Laboratory Animals* as well as the Brown University Institutional Animal Care and Use Committee (IACUC).

2.2. Stimuli

The visual stimuli were color bitmap images chosen from three categories of objects: baskets, insects, and pocket watches (Hemera Technologies Corporation; Seattle, Washington). Each category of objects comprised 16 individual objects, and each category could be divided in half by color (Fig. 1A). This yielded a total of 48 individual objects (3 categories, 16 objects in each category). In the experiments, the stimuli subtended approximately 6° of visual angle.

2.3. Apparatus

The animals were tested in isolated rooms sitting adjacent to the experimenters' workstations. Stimuli were displayed on a monitor placed 120 cm from the monkey, which covered approximately 19 × 14° of visual space. Manual response buttons were placed at arms' distance, just beyond two

openings in the front panel of the chair. The buttons were located approximately 35° visual angle below the bottom edge of the display, and were partially occluded by the primate chair, but not explicitly hidden from view. Experimental control and data collection of behavioral measures was conducted using custom written programs. This system uses computers running the QNX operation system (QNX Software Systems; Ottawa, Ontario) to provide deterministic control and acquisition of button responses and eye position and to communicate with a dedicated graphics machine using isolated high-speed Ethernet and direct digital I/O. For two monkeys, eye movements were recorded using an ISCAN RK-726PCI video eye tracking system, running at 120 Hz (ISCAN, Inc., Burlington, MA). For the third monkey, eye movements were monitored using both the ISCAN system and a Robinson style scleral eye coil system (CNC Engineering).

2.4. Procedure

The monkeys were first trained to sit in a primate chair in the separate isolated testing chamber and then to make button responses in a simple non-spatial color discrimination task. During this initial button training, the animals' heads were not restrained. After the initial training period, and following the implantation of the headpost, all experiments were conducted with the monkeys' head fixed and eye movements recorded.

Each monkey learned two basic tasks: a viewing only task and a visual classification task (Fig. 1). For both tasks, visual stimuli were presented in the center of the display monitor and subtended approximately 6° of visual angle. In the viewing only task (Fig. 1B), the monkeys initiated trials by fixating a yellow spot which subtended 0.3° of visual angle for 450 ms. Following acquisition, the fixation spot was extinguished and, following a 200 ms blank period, a visual stimulus was presented for 600 ms. At the end of the stimulus period, the screen was blanked for 250 ms and a second fixation spot was then presented in a randomly selected location 6° above, below, to the right or to the left of the center. The monkey was required to fixate the second spot to receive juice reinforcement.

In the classification task (Fig. 1C), the monkeys initiated trials by fixating a blue fixation spot for 450 ms presented in the center of the monitor. The fixation spot was removed from view once the fixation requirement was met and the test stimulus appeared following a 200 ms delay, as in the viewing only task. The test stimulus remained visible until the monkey made a button response or until 5000 ms had elapsed. The monkeys were given juice reinforcement for correct responses. The inter-trial interval was 1 s.

During training, the monkeys were shown 16 objects from the basket, insect, and pocket watch object categories, as described above. Two of the three categories were used in the recognition task, and the remaining category was reserved for the viewing only condition. The two object categories in the recognition task were learned at different levels of specificity. The stimuli used for the *color* version of the task were assigned their button response based on color (e.g., blue baskets were assigned to the left button and yellow baskets were assigned to right button). Stimuli in the *individual* category were assigned to each button randomly such that four objects of each color were assigned to each button (e.g., four blue baskets and four yellow baskets were assigned to the left button, and the remaining four in each color assigned to the right button). Thus, the monkeys could solve the classification task for objects in the *color* task using a simple color cue. In the *individual* task, however, the monkeys had to learn the correct association for each specific exemplar. Note that during all training and testing blocks, stimuli from the *color* and *individual* categories were intermixed. The third category of objects was shown in the viewing only task so that the monkeys were familiarized with the procedure of looking at stimuli without being required to press buttons. Recognition task training blocks consisted of 4 repetitions of each of 16 objects in two categories for a total of 128 trials in each block. Each monkey was required to complete a minimum of 100 blocks in the recognition task.

Following the recognition training, the monkeys were shown the images from the learned object categories in the viewing only condition. The prior training in the viewing only condition simplified this transition, and the monkeys had no problems withholding manual responses during this viewing task. Showing the well-known objects in the viewing task allowed us to test whether any pattern of eye movements that emerged during the

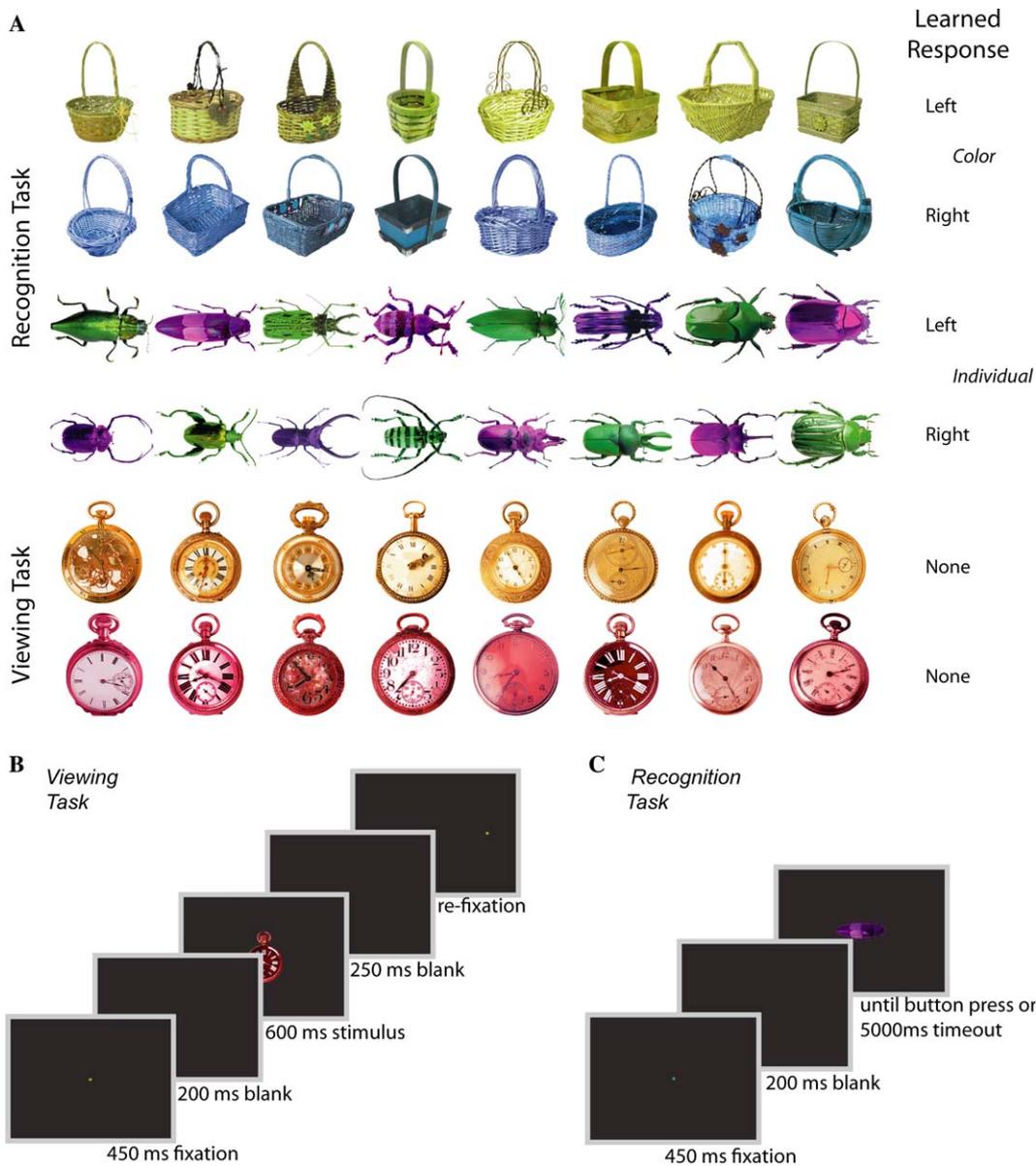


Fig. 1. Stimuli and task procedures. (A) A total of 48 images were used in the experiments, chosen from three separate categories (baskets, insects, and pocket watches). Individual categories contained exemplars of two different colors. A sample mapping of the stimuli to the tasks is shown. For this example, the baskets and insects would be assigned to the recognition task and the pocket watches would only be seen in the viewing only task. In addition, the insects would have to be learned individually, but the baskets could, in principle, be learned by color (press left for yellow baskets and right for blue baskets). Note that each of the monkeys in the experiments had a different mapping for stimuli and task. (B) Procedure used in the viewing task. For this task, the monkeys initiated the trial by fixating a spot at the center of the display for 450 ms. The spot was extinguished and replaced by an image 200 ms later. After 600 ms, the image was turned off and a new spot appeared randomly around the display. The monkeys were rewarded for looking to the location of the spot. (C) In the recognition task, the trials began as in the viewing only task, but the monkeys were required to press the assigned button for the displayed object within 5000 ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

recognition task would also be observed in a task in which the monkeys were not required to make any explicit recognition response. Viewing only blocks in this phase of the experiment consisted of two repetitions of each of the 32 objects (2 categories of 16 images) learned in the training phase. The monkeys completed 40 blocks in the post-training viewing only condition.

2.5. Eye movements

Eye movements were recorded throughout each trial of the experiment. Note that for all of the tasks in this study, there were no explicit constraints on the monkeys' eye movements during the time period that the visual stimuli were present on the display. The analog output from the eye tracking

hardware was sampled by the control system at 1 kHz and a moving average was stored to disk every 5 ms (200 Hz). Because the temporal resolution of the ISCAN system was limited to 120 Hz, great care was taken to ensure that the signals from this system were consistent with those from the eye coil system. We verified the timing relationship between the two trackers by recording analog eye movement signals from both systems simultaneously for one of the three subjects (Monkey J). Continuous recordings were then superimposed and scaled and offset to match, and we found the temporal offset which maximized the correlation between the two signals. The signal obtained by the camera system was found to be delayed by 16 ms (two 120 Hz video frames) with respect to the scleral search coil, so all data from the camera system were offset in time by this amount to account for the

delay. Saccades endpoints and timing extracted from both signals were highly consistent, and we observed no systematic biases or omissions in analyses based on the ISCAN data for this comparison.

Saccades were automatically extracted from offline eye records using a velocity-based algorithm written in C, which marked the start and end time, and start and end position for every saccade on each trial. The parameters of this algorithm were set to detect saccades down to approximately 0.4° in amplitude.

2.6. Data analysis

Behavioral performance during the recognition task was assessed as a function of both proportion correct trials and manual response times (the latency between the onset of the visual stimulus and the button press). Analysis of eye movement patterns during the recognition and fixation tasks focused on the initial saccade on each trial occurring between 50 and 350 ms following the onset of the visual stimulus. Only trials with initial saccades that occurred before the manual response and that were less than 3° in amplitude (within a circle surrounding the visual stimulus) were analyzed. Criteria for saccade analysis in the fixation task were the same, except that no manual response limits were imposed.

We used multivariate analysis of variance (MANOVA) with horizontal and vertical initial saccade endpoints as dependent variables to assess the effect of individual image, image category, and response on eye movements. To further quantify the relationship between observed saccade patterns and the learned manual responses in the recognition task we computed empirical “receiver (or relative) operating characteristic” (ROC) curves and estimated the area under these curves (Green & Swets, 1966; Swets, 1995). The technique employed here is similar to that used in previous physiological studies of saccade target selection (Horwitz & Newsome, 2001; Shadlen & Newsome, 1996; Thompson, Bichot, & Schall, 1997). In this experiment, the area under the ROC curve is an estimate of how well the eye movements predicted the individual response associations (left or right button press) that monkeys had learned during training. To compute the area under the ROC, we extracted the horizontal and vertical initial saccade endpoints for each trial, sorted by monkey and task (e.g., color, individual, and fixation). We then systematically stepped a criterion throughout the range of observed endpoint values and calculated the proportion of left button and right button trials that exceeded the criterion. The area under the ROC curve was calculated using the trapezoidal method. This value is restricted to the range between 0 and 1, with 0.5 being chance performance. The computed area can be interpreted as the probability of correct classification (in this case “left” or “right”) by an ideal observer using only the saccade parameter. The ROC values were calculated for both vertical and horizontal saccade endpoints. In addition, in some analyses we subselected data as a function of saccade latency in order to estimate whether the available information changed as a function of time during a trial.

Statistical significance for the area under the ROC curve was computed using a permutation test (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Horwitz & Newsome, 2001). For this test, we calculated the areas under ROC curves using the same trials with randomly permuted button assignments 1000 times, and found the exact probability of obtaining the area measure for the actual assignment if it came from the randomly assigned group.

3. Results

3.1. Task performance

Each monkey learned to classify visual images from two of the three sets of objects shown in Fig. 1. For one of the learned sets, the button mapping was based on the objects’ color (the *color* task). For the other set of objects, the button mapping was specifically not color based, so the stimuli

had to be learned individually (the *individual* task). The remaining set was only viewed during a fixation task. In this paper, we analyzed the monkeys’ steady state performance and therefore considered data from the second half of the training period (stimulus repetitions 200–400, blocks 50–100). For these trials ($n = 6400$ per monkey), percent correct performance was 99.2% (J), 98.4% (S), and 97.7% (T). Mean median reaction times (i.e., the average median response from each block of 128 trials) were 486 ms (J), 390 ms (S), and 397 ms (T).

3.2. Initial saccade endpoints depend on the visual stimulus

To understand how the monkeys’ eye movement patterns related to their performance in the recognition task, we first examined how frequently and at what time saccades occurred during the recognition task. We analyzed those trials which included at least one saccade prior to response and where the eye position remained within a $\pm 3^\circ$ window (within the bounds of the visual stimulus; see Section 2). This selection criterion left 98% (J), 94% (S), and 69% (T) of the trials for the following eye movement analysis. In general, Monkey T was much more likely than the other two animals to press a button before making a saccade (20% of 6400 trials vs. <1%, Monkey J and 2%, Monkey S). It is important to reiterate that the monkeys were neither directly rewarded nor punished for making eye movements, and it is thus not surprising that they adopted different oculomotor strategies.

As each monkey learned to recognize 32 different images in the classification task, we first asked whether initial saccades were in any way dependent on the image presented. Analysis of the initial landing position using a one-way MANOVA including horizontal and vertical endpoints as dependent variables and image as the independent factor revealed highly significant effects of image on both vertical and horizontal landing position (all p 's $\ll 0.01$). We repeated the analysis blocked by the animal’s actual response, to ensure that the effect of image was not solely attributable to the different manual responses. For these analyses, the image factor was still highly significant in all conditions. These results were by no means unexpected, as the test images were not all scaled identically and individual images had distinct external contours and internal features. They show, however, that the initial saccades metrics did depend on the image shown and were not simply a stereotyped, reflexive behavior.

3.3. Eye movements during recognition depend on learned responses

A detailed analysis of the initial saccades revealed a striking and consistent pattern of results for all three monkeys (Fig. 2). Each set of plots in Fig. 2 shows the endpoints of the initial saccades sorted by side of assigned response. These figures demonstrate that the initial saccade endpoints were highly predictive of an image’s learned button

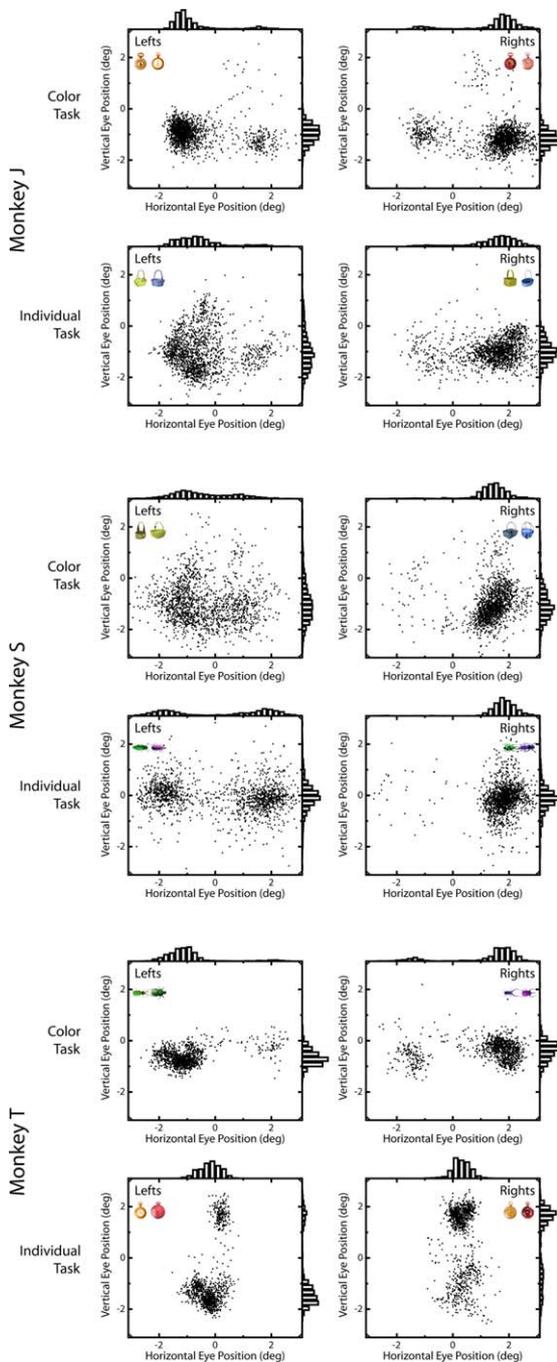


Fig. 2. Initial saccade endpoints predict side of manual response. Each plot contains markers denoting the endpoint of initial saccades for single trials in the recognition task. The six rows of plots are arranged with alternating *color* and *individual* conditions sorted by subject. The two columns separate trials in which a left button response and right button response was correct (plots labeled “lefts” and “rights”). Inset images show two stimuli from the particular condition. For example, the second row of plots shows trials in the *individual* task for Monkey J, wherein baskets were learned at the individual level (hence the yellow and blue baskets are present in both left and right columns). Saccade endpoints are placed at the corresponding horizontal and vertical position (in degrees), referenced to the center of the display. Histograms above and to the right of each plot show the marginal distributions in the horizontal and vertical dimensions, respectively. It is apparent that these distributions for the left and right conditions are well separated in all six conditions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

assignment. The eye movements clearly indicate that the visual images had, at some level, been recognized by the time the saccades were launched.

Using ROC analysis (see Section 2), we quantified how well the initial eye movement parameters predicted the side of response assigned to an image. In particular, we considered how the distributions of saccade endpoints could be used to determine the learned response for each trial. An example of this procedure is shown in Fig. 3. The figure shows how the distributions of horizontal saccades relate to the manual responses, as almost all leftward saccades occurred during trials containing left button objects, whereas rightward saccades were much more frequent when the visual stimulus required a right button press.

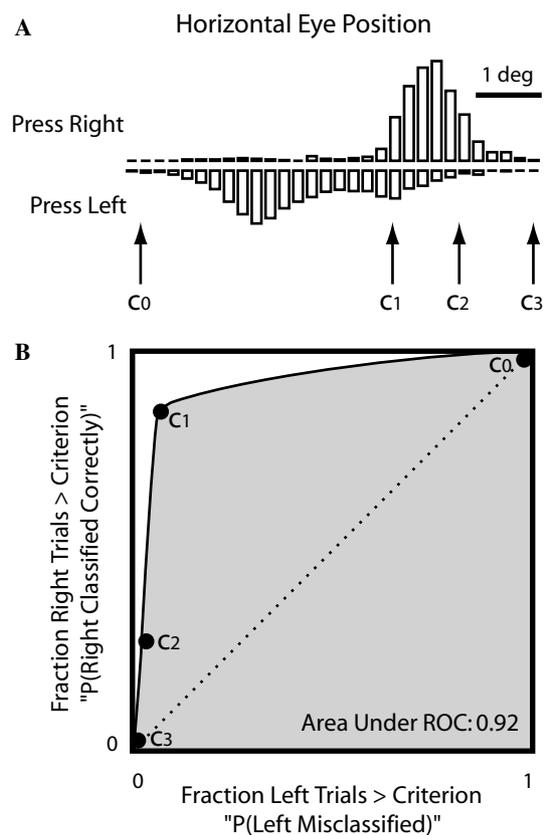


Fig. 3. ROC analysis method. (A) Horizontal saccade endpoint distributions from a single condition from Fig. 2 (Monkey S, *color* task) arranged with the right button trials on top (upward bars) and the left button trials on bottom (downward bars). A subset of possible threshold criteria (C0, C1, C2, and C3) are denoted by the arrows below. (B) From the criteria levels shown in (A), the fraction of left trials greater than the criteria (abscissa) is plotted against the fraction of right button trials greater than the criteria (ordinate). For example, for criterion C0, all left trials and right trials are greater than the criterion, resulting in the point (1, 1). Similarly, for criterion C3, no left trials or right trials are greater than the criterion, resulting in the point (0, 0). The closer that intermediate criteria results in points near (0, 1), the better the two distributions can be separated. The ability to separate the populations can be quantified by completing the curve between the two extreme criteria and calculating the area under this curve. Distributions that are completely overlapping will result in the line of slope one (dotted line) with an area under the ROC curve of 0.5. Perfect discriminability yields an area measure of 1.0.

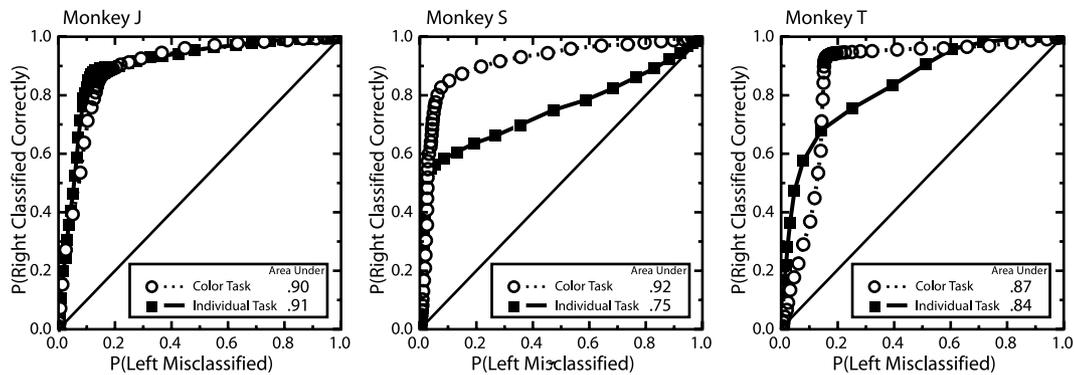


Fig. 4. ROC curves for the recognition conditions. Plots show the ROC curves for each monkey using the horizontal endpoint of the initial saccade, as in Fig. 3A. In each graph, the data from the monkey's *color* task are shown by the open circles and the *individual* task by the filled squares. Conventions are as in Fig. 3B. For all six conditions, the area under the ROC curve (shown in inset of each graph) was significantly greater than the chance level ($p < 0.01$, by permutation test), indicating that considerable information about an image's button mapping could be extracted from the horizontal landing position of the initial saccade.

In each of the six recognition conditions (three monkeys, two object categories each) the horizontal position of the first saccade provided a remarkably good predictor of assigned button response for the object presented on that trial (Fig. 4). Perfect classification performance as measured by the area under the ROC curve would be 1.0, with chance being 0.5. The computed areas under the ROC curves in Fig. 4 (insets) were all significantly different from chance ($p < 0.01$, permutation test).

We repeated this analysis using vertical, instead of horizontal, eye positions and in all six cases, the area under the ROC was less than that computed from the horizontal endpoints. This horizontal bias is clearly evident in Fig. 2, although one condition—Monkey T's individual task—seems better separated by the vertical eye movements. Even for this condition (Fig. 2, bottom plots) a direct comparison between the horizontal and vertical endpoint showed that despite the larger absolute vertical movements, the horizontal distributions are actually slightly better, in terms of the area under the ROC curve, in predicting the response side (area under curve 0.84 vs. 0.80). For simplicity, therefore, we used the horizontal endpoints in the subsequent ROC analyses.

As the horizontal component of the initial saccades for each of the monkeys was predictive of the associated manual responses, we were interested in how early, within a trial, this information became evident. Because the saccades did not all occur at precisely the same latency, we could bin trials by saccade time, and compute the area under the resulting ROC curves as a function of time following stimulus onset (Fig. 5). Sensitivity was measured for non-overlapping periods spanning the monkeys' saccade latency distributions, which are shown in the bottom portion of each of the graphs in Fig. 5. Monkey J showed no effect of saccade latency on the sensitivity measure, whereas Monkey S's sensitivity increased with increasing latencies, especially in the more visually demanding "individual" task. The sensitivity measure for Monkey T decreased slightly at the longest latency.

The cumulative saccade distribution for Monkey J (Fig. 5A, bottom) shows that this animal made more early saccades in the *color* condition compared to the *individual* condition, but saccades in both were highly predictive of the learned response (Fig. 5A, top). For Monkey S (Fig. 5B), the eye movements in the potentially simpler *color* task were slower than in the *individual* task (seen as a rightward shift in the cumulative distributions), but this delay resulted in higher overall sensitivity. This tradeoff is illustrated more clearly in Fig. 6, which shows the ROC curves for Monkey S sorted by task and by latency of eye movement following stimulus onset (before and after 180 ms). Interestingly, for the *individual* task objects (Fig. 6B), saccades launched before 180 ms were not sensitive to the objects' identities (area under ROC = 0.53, not significantly different from chance). Eye movements executed after 180 ms, however, were highly predictive (area under ROC = 0.86, $p < 0.001$ by permutation test). For the *color* task objects (Fig. 6A), both early and late saccades were predictive of the side of response (ROC area = 0.84 for early saccades, 0.92 for late saccades). This analysis suggests one potential approach for using detailed saccade related information as a temporal marker for recognition processes (see also, Kirchner & Thorpe, 2006).

3.4. Initial saccade latencies predict manual reaction times

The link between the saccade targets and the recognition response led us to examine the trial by trial correlation between the time of the first saccade and manual reaction time (Fig. 7). Between approximately 175 and 275 ms, a range which included almost all initial saccade latencies, there was a systematic increase in manual reaction times with increasing saccade latency.

A summary of the saccade latencies and manual reaction times for all three monkeys is shown in Fig. 8, and reveal a notable parallel between the initial saccade latencies and the manual reaction times. For example, Monkey T, who consistently responded more quickly with his left hand than

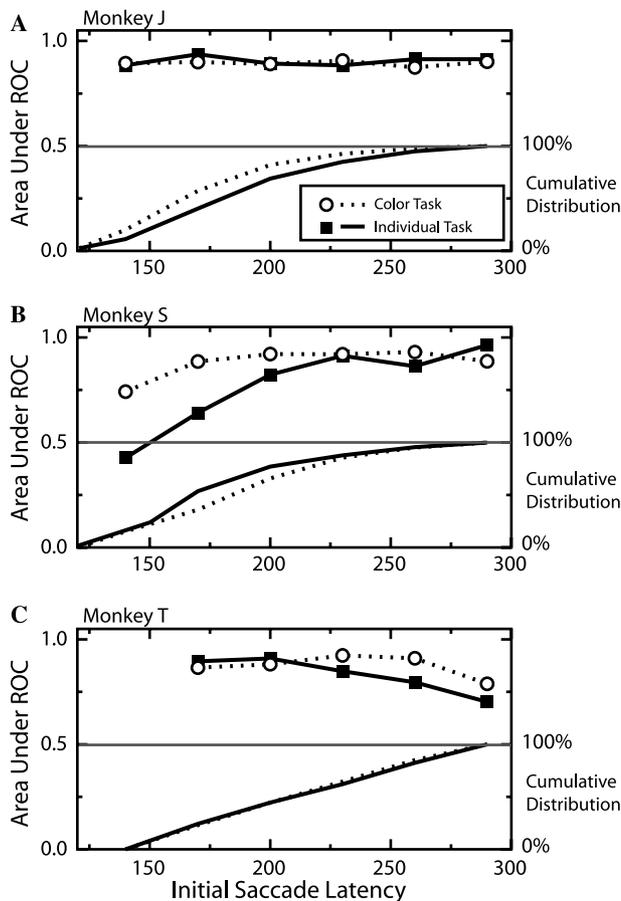


Fig. 5. Area under the ROC curve as a function of saccade latency for each of the three subjects (A–C). Trials were sorted by initial saccade latency (time from stimulus onset to the beginning of the first saccade) to determine if early or late saccades were more or less informative about an image's button mapping. Data from each monkey are presented in individual plots. Within a plot, the bottom halves of the split plot show the cumulative distributions for the percentage of saccades launched by a particular time. Data from the *color* task are shown by the broken lines and open circles, and the *individual* task data are marked by the solid lines and filled squares. Monkey J (A) showed little effect of saccade latency on the area measure for either the *color* or *individual* task, but the cumulative distribution of saccade times indicates that initial saccades in the *color* task were launched earlier. Monkey S's data (B) indicate that longer latency saccades were more informative, as the curves for both tasks increase over time, with the sharpest increase seen for the *individual* task. Monkey T's initial saccade latencies (C) were slowest overall and no benefit (in terms of area under the ROC) for longer latency saccades was observed.

with his right also consistently made faster initial saccades for images associated with the left button. Analogously, Monkey S, who responded faster with his right hand than his left, made faster initial saccades for trials containing right button targets.

3.5. Impact of "incorrect" initial saccades

Although we found that even in the absence of any explicit oculomotor instruction, the monkeys' initial saccades were robust predictors of the learned hand-image associations, they were not perfect predictors. We could

therefore ask whether there was any effect of correct or incorrect initial saccade direction on the monkeys' manual responses. For this analysis, we first determined the criterion level for horizontal position that led to the highest accuracy, i.e., the criterion an ideal observer should use in order to correctly classify the largest proportion of both left and right target objects using only the eye position information. With these criterion set for each monkey and each task (color or individual), we labeled initial saccades "correct" if they correctly classified the response side (left or right) for a stimulus. Paired *t*-tests for each monkey showed that accuracy decreased reliably on trials where the initial saccade incorrectly predicted response side. The effect on reaction times was even clearer. Fig. 9 shows the mean manual response times for correctly identified targets sorted by monkey, task, and whether the initial saccade accurately predicted the response. Taken together, the effect of initial eye movement (correct vs. incorrect) was highly significant (mean 448 ms vs. 551 ms). Thus, on trials where the initial eye movements misclassified the trial type, the monkeys' manual responses were both less accurate and significantly slower.

3.6. Eye movements during passive viewing

In the classification task, the process of recognizing an object (i.e., recognizing it as a "left" or "right" button object) cannot be clearly dissociated from the physical action of making the associated response (move hand to press left button or right button). Would the oculomotor pattern be evident even in the absence of the overt motor act? To address this question, we first familiarized the monkeys with the process of "passively" viewing visual objects by using the third set of objects not used in the recognition task. For this task, button presses were not allowed (and would have aborted the trial, had they occurred). Following the recognition training with two of the three stimulus sets, we showed these objects in the viewing only task. As in the above analysis, we extracted eye movements for each trial and created distributions of initial saccade position following the onset of the visual stimulus. For these data, there was no actual manual response. We could still ask, however, how well the eye movements predicted the previously associated side of response. The ROC curves from these data are shown in Fig. 10. For all three monkeys, the pattern of initial eye movements carried information about which button the images had been associated with, even though during the viewing task no buttons were actually pressed. There was variability in the predictiveness of these eye movements, as Monkey J showed almost identical eye movement patterns in this task as in the recognition task, whereas the saccade patterns for Monkeys S and T were less informative than in the active recognition task. We can say, though, that neither the physical act of pressing the button nor the active requirement to press buttons is necessary to evoke a learned pattern of oculomotor movements. Both action and intention, however, likely contribute to

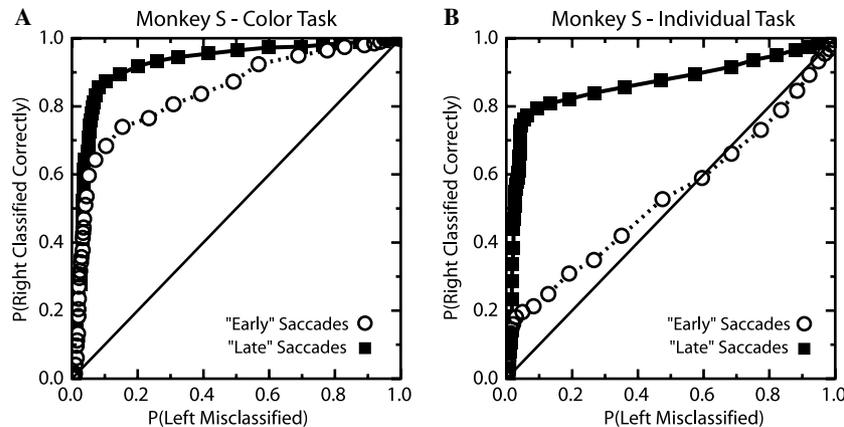


Fig. 6. Saccade latency as a probe for the speed of visual recognition in different tasks. ROC curves for Monkey S separated by *color* (A) and *individual* (B) task and by saccade latency. Trials were split into two groups as a function of saccade latency. “Early” trials were those in which the initial saccade occurred before 180 ms and the “Late” trials were those in which the initial saccade occurred at or after 180 ms. For the color task (A), both early trials and late trials were informative about an image’s button mapping, with a slight advantage for the potentially simpler *color* task (area under ROC for *color* task: 0.84, *individual* task: 0.92). For the *individual* task (B), no significant information about an image’s button mapping was evident for early saccades (area under ROC, 0.53). Trials with later saccades, though, showed a large increase in information about the image mapping (area under ROC, 0.86).

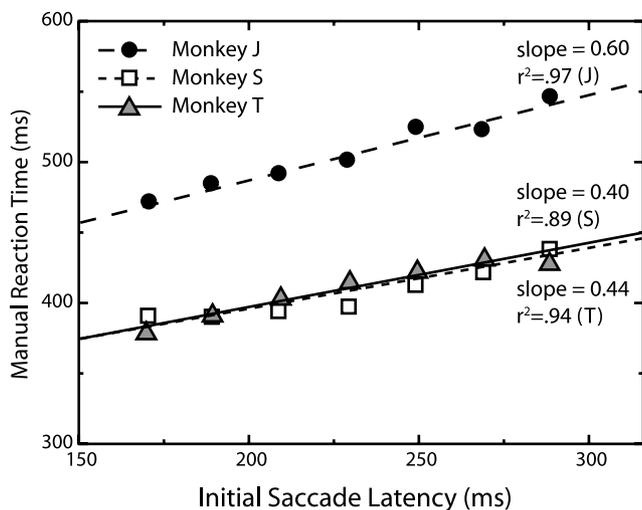


Fig. 7. Correlation between manual reaction times and initial saccade latencies, suggesting a possible link between the two actions. Trials were binned according to saccade latency into seven groups (range: 160–300 ms, bin size: 20 ms). Median manual response times for each group were then calculated for each of the three monkeys and then plotted against the midpoint of the latency for that bin. Each monkey showed a systematic increase in reaction time for increasing initial saccade latency.

more precise linkages between the eye and hand, and different animals may take different strategies when the same stimuli are employed in two different tasks.

4. Discussion

We found that monkeys’ eye movement patterns during a visual recognition task reflect learned recognition responses, which for our task were manual button presses. Interestingly, the stimulus image affected both the direction of the saccade, which was related to learned response associations, and the specific landing position, which depended on the particular image features.

What is the nature of the linkage between the oculomotor and manual responses? Because the initial saccade targets were so clearly tied to the response side associated with the images, the saccade metrics do not seem to carry significant information about what image information was actually used by the monkeys to classify the images. Thus, for our task, saccade endpoints do not seem to reflect the active selection of informative regions. Instead, the endpoint of the initial saccade was related to the learned response of the visual stimuli, implying that at the time the eye movements were executed, adequate information had already been extracted to activate this learned association. In other words, properties of the first eye movement tell us that a large component of the recognition process has already occurred.

One explanation for the tight coupling between the oculomotor and manual responses is that these processes share a common decision path. This view is supported by data showing that programming and execution of eye, head, and arm movements are more coordinated than overt response latencies might suggest (Biguer, Jeannerod, & Prablanc, 1982). In our experiment, the linkage between the oculomotor and manual responses was also evident in the reaction time correlations shown in Fig. 7, which are consistent with previous human studies that have reported modest, but significant, trial to trial correlations between saccade and limb movement latencies (Bekkering, Adam, Kingma, Huson, & Whiting, 1994; Bekkering, Adam, van den Aarsen, Kingma, & Whiting, 1995; Frens & Erkelens, 1991).

It is somewhat surprising that on the large majority of trials, the monkeys consistently produced short amplitude saccades between 2° and 3° before the manual response. In most conditions we also observed a slightly downward bias. These metrics suggest a parallel between the spatial attentional mechanisms involved in selecting a response location and simultaneously redirecting gaze. At the same time,

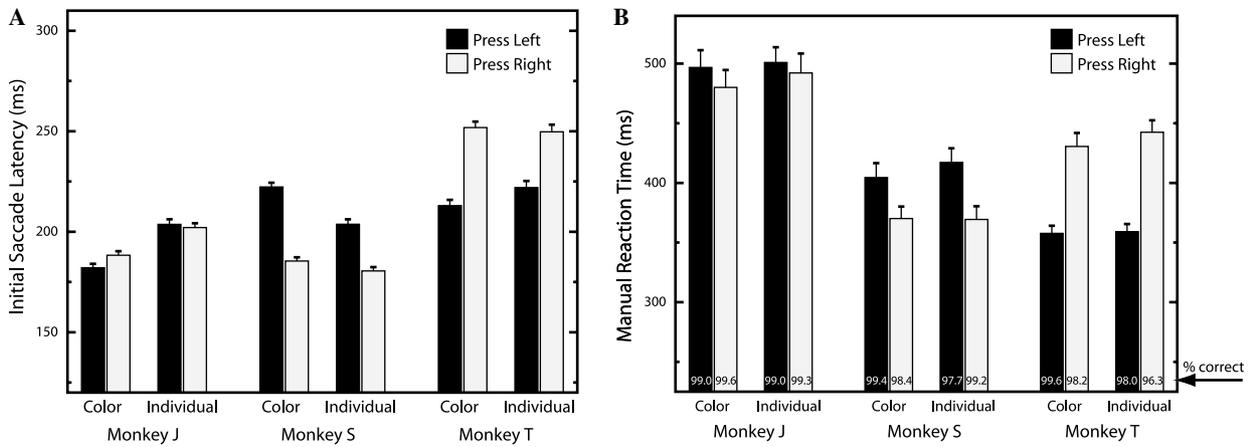


Fig. 8. Links between initial saccade latencies (A) and manual reaction times (B) in the recognition tasks. In (A), mean saccade latencies are shown for each monkey in their respective *color* and *individual* tasks, and are further divided by the side of response for the image presented (dark and light bars). In (B), the manual response times are sorted in the same way, and a comparison of the two shows that the correlations seen in Fig. 7 are further evident in side biases. Monkey S shows speeded right button presses and rightward saccades, whereas Monkey T has speeded left movements for both the hand and eye movements. Monkey J is relatively unbiased for both, but does show a significant slowing in saccade latencies for the *individual*, compared to the *color*, task. Error bars denote 95% confidence intervals.

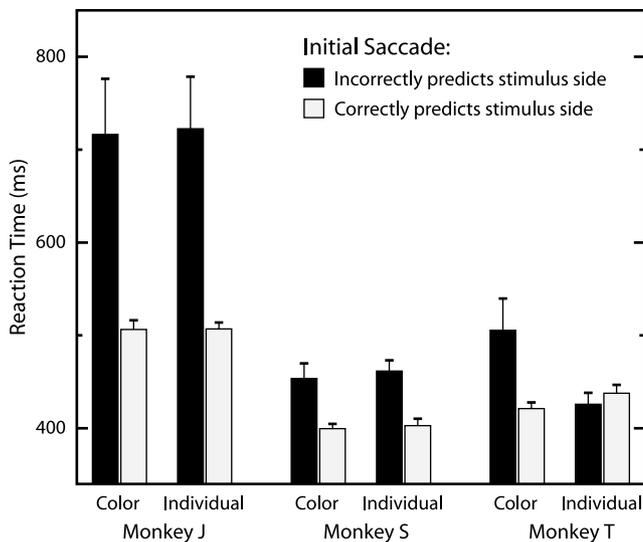


Fig. 9. Effect of initial saccade direction on manual reaction times. Plot organization is similar to Fig. 8B, except that trials in each of the six conditions are sorted by whether the initial saccade was in the “correct” direction for predicting the assigned button (see text for details). Reaction times for correct side trials are significantly speeded compared to incorrect saccade trials for five of the six conditions. Note that accuracy for all of the conditions exceeded 94% correct. Monkey J was particularly slowed on trials in which the initial saccade direction incorrectly predicted the correct manual response. Error bars denote 95% confidence intervals.

though, the scale of the one visible object (approximately 6°) also constrained the metrics of this initial saccade, as almost all initial saccades landed on a portion of the visual object. It thus seems that the actions associated with the visual objects (press a button in a particular spatial location) systematically and predictably biased the target selection process. In the course of normal visuomotor activity, such a link is also likely to occur. Previous studies of the relationship between the eyes and hands during natural tasks (Hayhoe et al., 2003; Land et al., 1999), show that

during normal activities the eyes are regularly redirected to locations of impending object manipulations. Furthermore, Tipper et al. (2001) found that eye movements to a target are affected by whether or not a concurrent reach to the same target is required.

Many previous studies have used eye movements as an operant response. Here, we show that even under conditions in which saccades are not specifically rewarded, these movements can provide evidence regarding the time course of visual recognition. We found that under some circumstances it was possible to use the eye movement latency distributions to determine how quickly, within a trial, an object’s identity could influence the initial saccade direction. The data from Monkey S most clearly addressed this question, as the impact of the visual target’s identity increased quickly between 150 and 200 ms. Our data thus suggest one method for observing the rate at which sensory evidence in various tasks accrues and is transformed into motor responses, although the generality of this method is not yet certain. These findings are consistent with recent physiological studies probing the evolution of eye movement programming (Bichot, Chenthal Rao, & Schall, 2001; Gold & Shadlen, 2003). Gold and Shadlen (2003), for example, found that motor plans for executing overt responses are influenced throughout the period of sensory analysis. By interrupting the decision process with an electrically evoked eye movement, they could probe the buildup of decision related bias in the oculomotor response, thus demonstrating that there is a constant interplay between putative perceptual and motor systems. We found that eye movements naturally initiated in the course of a visual recognition task also reflect actions associated with the sensory stimulus being processed.

As the average initial saccade latency in the classification tasks was about 200 ms, the visual information used to bias

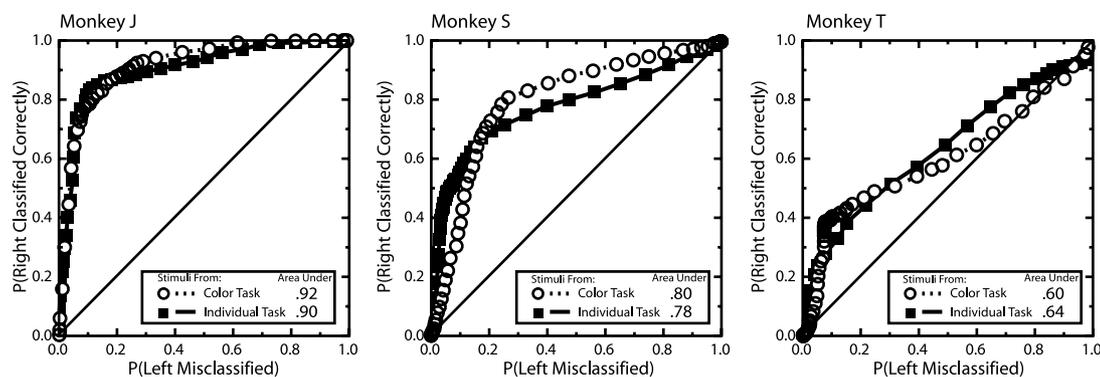


Fig. 10. Eye movements during “passive” viewing predict side of associated response for learned stimuli. Graphs are organized as in Fig. 4, except that data are taken from trials during which no explicit manual response was required (or allowed). During this viewing only condition, initial saccades continued to carry information about the learned button mappings, as areas under each of the curves was significantly different from chance. Areas under the curve for each condition are given in the inset for each plot and can be compared to the values shown in Fig. 4 for the full recognition trials.

these eye movements must have arrived sometime before the eye movements themselves. The complexity of the relatively large stimulus battery used in the current experiments suggests that extrastriate visual areas of the ventral pathway are likely recruited to solve the required discriminations. Given that the latency of visual neurons in the cortical areas of the temporal lobes thought to be crucial in complex visual discriminations is 100 ms or more (Desimone, Albright, Gross, & Bruce, 1984; Perrett et al., 1984; Sheinberg & Logothetis, 2001), there is less than 100 ms between the time these cells respond and the onset of the saccades whose direction is affected by their actions. This would suggest that the information about stimulus identity, theoretically available in the initial few spikes of neuronal activity (Tovee, Rolls, Treves, & Bellis, 1993), is functionally critical for recognition. This view is supported by recent theoretical models of rapid visual processing (Thorpe et al., 2001) and detection tasks using eye movements as an overt response mode (Kirchner & Thorpe, 2006).

Following an initial period of recognition training, we examined whether the observed pattern of eye movements would be evident even during a viewing only task, where no manual response was required. For all three animals, we found that initial saccades were still biased in the direction of the learned manual responses. The implication is that when known objects are viewed, this activity is never purely passive, and that eye movements can provide overt evidence for automatic recognition processes.

Intimately bound to both perceptual and motor processes, the oculomotor system highlights the natural interplay between sensation and action. How we look at objects in the world is linked to how we have learned to act on those objects. We have no evidence that the subjects in this experiment were aware of their eye movement patterns, nor do we know how interruptions in this oculomotor behavior might affect subsequent recognition judgments. Even so, the systematic oculomotor bias we have observed reveals how quickly perceptual recognition affects multiple action preparation systems in the brain.

Acknowledgments

This work was supported by the James S McDonnell Foundation, NSF CRCNS 0423031, and NIH RO1-EY014681.

References

- Bekkering, H., Adam, J. J., Kingma, H., Huson, A., & Whiting, H. T. (1994). Reaction time latencies of eye and hand movements in single- and dual-task conditions. *Experimental Brain Research*, *97*(3), 471–476.
- Bekkering, H., Adam, J. J., van den Aarsen, A., Kingma, H., & Whiting, H. T. (1995). Interference between saccadic eye and goal-directed hand movements. *Experimental Brain Research*, *106*(3), 475–484.
- Bichot, N. P., Chenthal Rao, S., & Schall, J. D. (2001). Continuous processing in macaque frontal cortex during visual search. *Neuropsychologia*, *39*(9), 972–982.
- Biguer, B., Jeannerod, M., & Prablanc, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. *Experimental Brain Research*, *46*(2), 301–304.
- Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, *13*(1), 87–100.
- Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Carpenter, R. H. S. (1988). *Movements of the eyes* (2nd ed.). London: Pion.
- Cutting, J. E., Alliprandini, P. M., & Wang, R. F. (2000). Seeking one’s heading through eye movements. *Psychonomic Bulletin & Review*, *7*(3), 490–498.
- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*(8), 2051–2062.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837.
- Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. New York: Oxford University Press.
- Frens, M. A., & Erkelens, C. J. (1991). Coordination of hand movements and saccades: evidence for a common and a separate pathway. *Experimental Brain Research*, *85*(3), 682–690.
- Gold, J. I., & Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *Journal of Neuroscience*, *23*(2), 632–651.
- Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

- Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, 3(1), 49–63.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7(11), 498–504.
- Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 210–228.
- Horwitz, G. D., & Newsome, W. T. (2001). Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. *Journal of Neurophysiology*, 86(5), 2543–2558.
- Judge, S. J., Richmond, B. J., & Chu, F. C. (1980). Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Research*, 20(6), 535–538.
- Kinsbourne, M. (1972). Eye and head turning indicates cerebral lateralization. *Science*, 176(34), 539–541.
- Kinsbourne, M. (1974). Direction of gaze and distribution of cerebral thought processes. *Neuropsychologia*, 12(2), 279–281.
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Research*, 46(11), 1762–1776.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35(13), 1897–1916.
- Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience*, 3(12), 1340–1345.
- Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, 28(11), 1311–1328.
- Land, M. F., & Tatler, B. W. (2001). Steering with the head: the visual strategy of a racing driver. *Current Biology*, 11(15), 1215–1220.
- Neggers, S. F., & Bekkering, H. (2002). Coordinated control of eye and hand movements in dynamic reaching. *Human Movement Science*, 21(3), 349–376.
- Noton, D., & Stark, L. W. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11, 929–942.
- Oliva, A., Torralba, A., Castelano, M. S., & Henderson, J. M. (2003). Top down control of visual attention in object detection. *IEEE Proceedings of the International Conference on Image Processing*, 1, 253–256.
- Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., et al. (1984). Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, 3(4), 197–208.
- Prablanc, C., Echallier, J. F., Komilis, E., & Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target. I. Spatio-temporal characteristics of eye and hand movements and their relationships when varying the amount of visual information. *Biological Cybernetics*, 35(2), 113–124.
- Remington, R. W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 6(4), 726–744.
- Robinson, D. A. (1963). A Method of Measuring Eye Movement Using a Scleral Search Coil in a Magnetic Field. *IEEE Transactions on Biomedical Engineering*, 10, 137–145.
- Robinson, D. A. (1981). The use of control systems analysis in the neurophysiology of eye movements. *Annual Review of Neuroscience*, 4, 463–503.
- Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: seeing and deciding. *Proceedings of National Academy of Sciences of the United States of America*, 93(2), 628–633.
- Sheinberg, D. L., & Logothetis, N. K. (2001). Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. *Journal of Neuroscience*, 21(4), 1340–1350.
- Swets, J. (1995). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc..
- Thompson, K. G., Bichot, N. P., & Schall, J. D. (1997). Dissociation of visual discrimination from saccade programming in macaque frontal eye field. *Journal of Neurophysiology*, 77(2), 1046–1050.
- Thorpe, S., Delorme, A., & Van Rullen, R. (2001). Spike-based strategies for rapid processing. *Neural Networks*, 14(6–7), 715–725.
- Tinker, M. A. (1946). The study of eye movements in reading. *Psychological Bulletin*, 43, 93–120.
- Tipper, S. P., Howard, L. A., & Paul, M. A. (2001). Reaching affects saccade trajectories. *Experimental Brain Research*, 136(2), 241–249.
- Tovee, M. J., Rolls, E. T., Treves, A., & Bellis, R. P. (1993). Information encoding and the responses of single neurons in the primate temporal visual cortex. *Journal of Neurophysiology*, 70(2), 640–654.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.