

Active inhibition and memory promote exploration and search of natural scenes

Paul M. Bays

Sobell Department of Motor Neuroscience and
Movement Disorders, UCL Institute of Neurology,
Queen Square, London, UK



Masud Husain

UCL Institute of Cognitive Neuroscience,
Queen Square, London, UK



Active exploration of the visual world depends on sequential shifts of gaze that bring prioritized regions of a scene into central vision. The efficiency of this system is commonly attributed to a mechanism of “inhibition of return” (IOR) that discourages re-examination of previously-visited locations. Such a process is fundamental to computational models of attentional selection and paralleled by neurophysiological observations of inhibition of target-related activity in visuomotor areas. However, studies examining eye movements in naturalistic visual scenes appear to contradict the hypothesis that IOR promotes exploration. Instead, these reports reveal a surprisingly strong tendency to shift gaze back to the previously fixated location, suggesting that refixations might even be facilitated under natural conditions. Here we resolve this apparent contradiction, based on a probabilistic analysis of gaze patterns recorded during both free-viewing and search of naturalistic scenes. By simulating saccadic selection based on instantaneous influences alone, we show that the observed frequency of return saccades is in fact substantially less than predicted for a memoryless system, demonstrating that refixation is actively inhibited under natural viewing conditions. Furthermore, these observations reveal that gaze history significantly influences the way in which natural scenes are explored, contrary to accounts that suggest visual search has no memory.

Keywords: eye movement, saccade selection, inhibition of return, memory, natural scenes

Citation: Bays, P., & Husain, M. (2012). Active inhibition and memory promote exploration and search of natural scenes. *Journal of Vision* 12(8):8, 1–18, <http://www.journalofvision.org/content/12/8/8>, doi:10.1167/12.8.8.

Introduction

Our ability to efficiently extract information from the visual world relies on rapid eye movements that bring different parts of the image into the fovea. These fixations are not random but selectively targeted towards locations of interest (Buswell, 1935; Henderson, 2003; Yarbus, 1967). Computational models (Itti & Koch, 2001) propose that target selection is based on an internal representation of the attentional priority assigned to each location in the scene.

The strength of activation within this “priority map” may be determined by a combination of the inherent saliency of visual elements at a given location, and their relevance to current behavioral goals. Activity consistent with such a priority representation has been identified in several areas of the oculomotor network (Fecteau & Munoz, 2006; Gottlieb, Kusunoki, & Goldberg, 1998; McPeck & Keller, 2002; Schall & Thompson, 1999; Thompson & Bichot, 2005).

According to theoretical accounts, competition within the priority map results in an eye movement to the location with the highest activity. So that gaze does

not repeatedly return to the location with the greatest attentional priority, an essential component of current models is a mechanism of “inhibition of return” (IOR) that discourages re-selection of previously-fixated locations (Klein, 1988; Klein & MacInnes, 1999; Posner & Cohen, 1984). Such a mechanism is supported by observations of reduced or delayed responses to previously-examined targets in brain regions associated with priority representation (Bichot & Schall, 2002; Bisley and Goldberg, 2003; Dorris, Klein, Everling, & Munoz, 2002; Fecteau & Munoz, 2011; Mirpour, Arcizet, Ong, & Bisley, 2009).

However, as a behavioral phenomenon affecting eye movements, the evidence for IOR is based almost exclusively on observations of saccade *latency*: fixations that precede an eye movement back to a previous location tend to be of longer than average duration (Farrell, Ludwig, Ellis, & Gilchrist, 2010; Hooge & Frens, 2000; Reuter-Lorenz, Jha, & Rosenquist, 1996; Vaughan, 1984). By contrast, studies examining the *frequency* of return saccades in naturalistic scenes seem to contradict the hypothesis that IOR acts as a “foraging facilitator” (Klein, 2000). These reports instead reveal a strong tendency to shift gaze back to

the location of the previous fixation (Hooge, Over, van Wezel, & Frens, 2005; Smith & Henderson, 2009, 2011a, 2011b). This surprising result appears incompatible with a spatial IOR, and has even led to the proposal of a “facilitation of return” that actively encourages refixations.

To determine whether return saccades are inhibited or facilitated in natural viewing, it is not sufficient simply to count their frequency relative to eye movements to other locations. This is because instantaneous influences on saccade selection (i.e., factors that are independent of where the eyes have been before) may indirectly affect the frequency of return saccades. Here we identify two classes of instantaneous influence on saccade selection—*attentional* biases that lead observers to preferentially fixate certain locations within an image, and *oculomotor* biases that make particular amplitudes or directions of eye movement more likely—and show that each independently contributes to the prevalence of return saccades.

To assess whether there is a true inhibition or facilitation of return saccades, therefore, we must compare their observed frequency with a baseline frequency corresponding to their expected prevalence if gaze history were irrelevant. If gaze planning was memoryless (i.e., a Markov process), saccade selection would be fully described by the conditional probability of choosing any target location within an image *given the current gaze position*, regardless of fixation history (Gardiner, 2010). Here we extract this conditional probability distribution from eye movements recorded in two naturalistic conditions: free-viewing of natural scenes, and search among everyday objects.

Return saccades, while common, were in fact substantially less frequent than would be expected if selection and search were memoryless, as proposed by some accounts (Horowitz & Wolfe, 1998). Instead, our findings reveal that exploration and search of natural scenes are indeed associated with active inhibition of return to previously-fixated locations.

Methods

Experimental procedure

Twenty subjects (nine male, 11 female; age 19–67) participated in the study after giving informed consent. All subjects reported normal color vision and had normal or corrected-to-normal visual acuity. Subjects sat with their head supported by a forehead- and chin- rest. Images were displayed on a 24” widescreen TFT monitor, viewed at a distance of 60 cm, covering an area of the visual field 30° high and 46° across. Eye position was monitored online at 1000 Hz using an

infra-red eye tracker (SR Research, Ontario, Canada). Each subject participated in two different conditions presented in counterbalanced order: *free-viewing* and *search*.

In the *free-viewing* condition, subjects were sequentially presented with twenty full-color photographs of natural scenes, comprising a mixture of people, wildlife, man-made objects and natural landscapes. Each trial began with a fixation cross in the centre of a blank display: once a stable fixation was recorded on the cross, an image was presented and gaze position recorded for twenty seconds, followed by a three second blank period before the start of the next trial. Every subject viewed the same images in the same order. The only instruction was to “look at the pictures.”

In the *search* condition, subjects were presented with twenty images each comprising a haphazard arrangement of forty everyday objects that one might find on a cluttered desk (e.g., a pen, a cup, a key, a coin). At the start of each trial, one object – the search target – was presented in isolation in a canonical orientation. This was the item that participants had to search for on that trial. It was followed by a central fixation cross as above; then the image was presented and gaze position recorded until the subject stopped the trial with a button press. Subjects were instructed to search the image for the target object and press the button either as soon as the target was located, or once they were certain it wasn’t present in the scene; they then indicated their response verbally.

The sequence of targets and images was the same for every subject, and the target item was present in 50% of images. To ensure all eye movements were representative of active search, only trials on which the target was *absent* were included in the analysis, thereby excluding any artifact associated with successfully locating the target.

Initial eye position analysis

The raw gaze position signal from the eye tracker was first low-pass filtered at 50 Hz to remove high-frequency noise (zero-lag five-sample Butterworth filter), and a gaze velocity estimate obtained by a two-sample difference filter. Blinks and periods where gaze was directed outside the display area were identified and removed. Fixations were identified by a combined velocity and duration criterion: each period of low gaze velocity ($< 50^\circ \text{ s}^{-1}$) lasting 100 ms or longer was identified as a fixation, and the mean gaze position during this period determined the fixation location. In total ~23,000 fixations were recorded in the *free-viewing* condition (mean 56 fixations per subject per

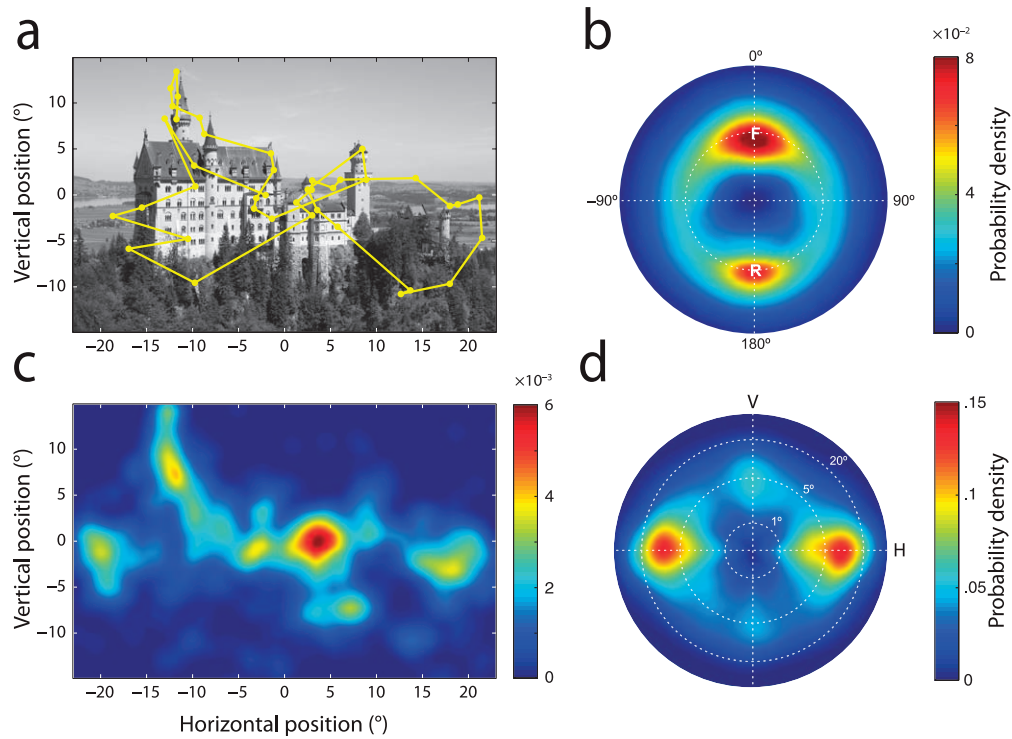


Figure 1. Gaze statistics in free-viewing of natural scenes. **(a)** One of the experimental images, overlaid with a typical *scanpath* (sequence of fixations and saccades) recorded during free-viewing. Images were full-color but shown here in grayscale for clarity. **(b)** Heat map on a polar plot showing the distribution of *relative saccade metrics* in free-viewing data. Hotter colors indicate greater probability density. Position on the circle indicates the angular difference between two successive eye movements. Here, 0° corresponds to a saccade in the *same* direction as the one before, i.e., the ‘forward’ direction. Positive values indicate clockwise deviation. 180° corresponds to the ‘return’ direction. Radial position indicates the difference in amplitude between successive saccades on a logarithmic scale: the dotted circle indicates equality. Note the two peaks corresponding to *forward saccades* (marked ‘F’) and *return saccades* (marked ‘R’). **(c)** Probability of fixation as a function of location (*fixation density*) for all participants viewing the image shown in (a). **(d)** Probability distribution of saccade amplitude and direction (*saccade density*) for all participants and images in the free-viewing task shown on a polar plot. Here V corresponds to vertical movements and H to horizontal ones. Radial position indicates saccadic amplitude on a logarithmic scale. Note how most saccades are horizontal and of relatively small amplitude.

image) and $\sim 10,000$ in the *search* task (mean 25 per subject per image).

Relative saccade metrics

To examine the effect of previous eye movements on saccade selection, we calculated the difference between each successive pair of saccade metrics in our recorded data,

$$\mathbf{m}_i = \mathbf{s}_i - \mathbf{s}_{i-1} = \begin{pmatrix} \Delta\theta_i \\ \Delta r_i \end{pmatrix}, \quad (1)$$

where $\Delta\theta$ indicates the angular difference between one saccade and the next (positive values indicating a clockwise deviation) and Δr indicates the log ratio of the saccade amplitudes (positive values indicating that the second eye movement is larger).

A *forward saccade* (i.e., an eye movement of exactly the same amplitude and direction as the one preceding it) would therefore correspond to $\mathbf{m}_i = \begin{pmatrix} 0^\circ \\ 0 \end{pmatrix}$ and a *return saccade* (i.e., an eye movement resulting in an accurate refixation) would take the value $\begin{pmatrix} \pm 180^\circ \\ 0 \end{pmatrix}$. The distribution of relative saccade metrics observed in free-viewing data is shown in [Figure 1b](#).

Estimating fixation density

A standard approach to analysis of gaze data is to treat the set of fixation locations recorded on each image, pooled across subjects, as independent and identically distributed samples from an unknown population distribution (the ‘‘fixation density’’). Kernel density estimation (e.g., Silverman, 1986) is a common

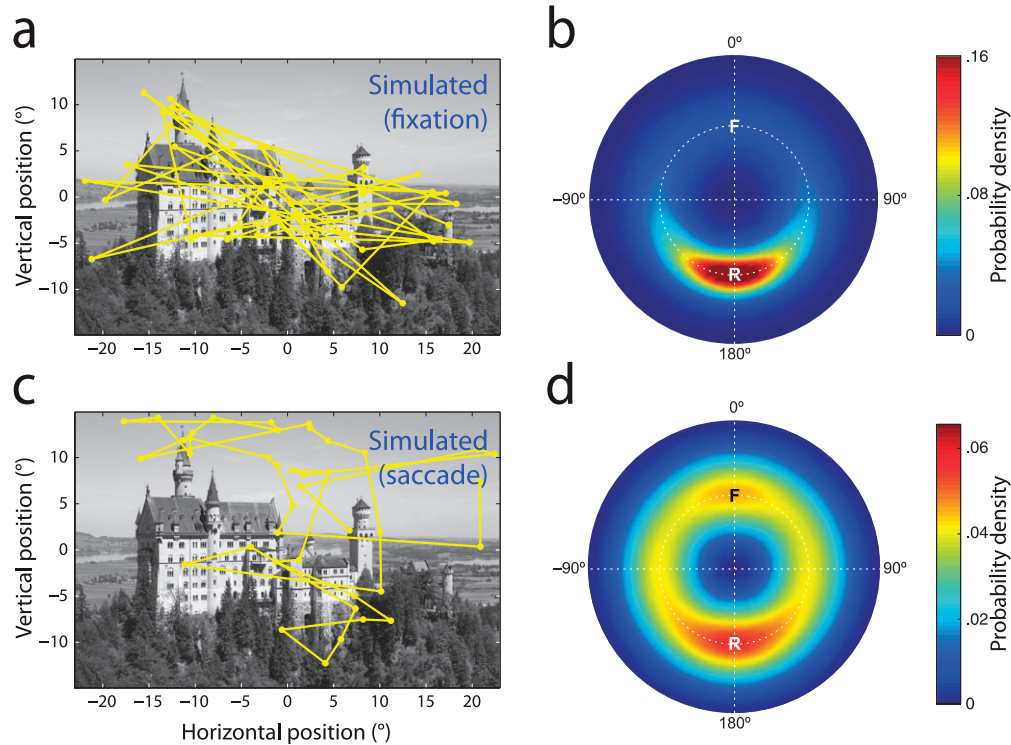


Figure 2. Attentional and oculomotor biases contribute to return frequency. **(a)** Typical scanpath generated by random sampling from the *fixation density* estimated for an example image (Figure 1c). **(b)** Relative saccade metrics in a large set of simulated scanpaths based on sampling fixation density. Note prevalence of *return saccades* (R), despite *independence* in selection of each simulated fixation. **(c)** Typical scanpath generated by random sampling from the *saccade density* (Figure 1d) estimated for free-viewing data. **(d)** Relative saccade metrics in simulated scanpaths based on saccade density. Note prevalence of both *return* (R) and *forward saccades* (F), despite independence in selection of each simulated saccade.

method of obtaining such a distribution: given a finite sample $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ of fixations at Cartesian coordinates $\mathbf{x}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$, the underlying probability density that a fixation \mathbf{x}_i falls at location \mathbf{X} is estimated by:

$$\hat{p}(\mathbf{x}_i = \mathbf{X}) = \frac{1}{n} \sum_i K_h(\mathbf{X} - \mathbf{x}_i), \quad (2)$$

where K is a two-dimensional Gaussian ‘kernel’:

$$K_h \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{h^2} \phi(x/h) \phi(y/h), \quad (3)$$

x and y are horizontal and vertical components of the input, ϕ is the standard Gaussian function with mean zero and variance one, and h is the kernel ‘bandwidth’. Rather than hand-pick the bandwidth, we used log-likelihood cross-validation (Habbema, Hermans, & Van Den Broek, 1974) to select the optimal bandwidth for each data set:

$$h = \arg \max_h \sum_I \left[\sum_i \log \sum_{j \neq i} K_h(\mathbf{x}_i^I - \mathbf{x}_j^I) \right], \quad (4)$$

where \mathbf{x}_i^I indicates the i th fixation on the I th image. This provided a principled method of obtaining a fixed kernel bandwidth for each dataset (free-viewing: $h = 0.93^\circ$; search: $h = 1.14^\circ$) that would provide the best description of the data. Figure 1c shows the resulting estimate of fixation density for an example image.

To examine how preferences for different regions of an image affect gaze statistics, we performed a Monte Carlo simulation based on the estimated fixation densities. The simulation proceeded as follows: 2,000 sequences of fixations, each matched in length to the experimental data, were generated for each image based on pseudorandom sampling from the fixation distribution estimated in Equation 2. Relative saccade metrics were calculated for these simulated scanpaths (example in Figure 2a) in the same way as for recorded data (Equation 1), and frequencies averaged across simulation repetitions to obtain the distribution in Figure 2b.

Estimating saccade density

The probability of generating eye movements of *different sizes and directions* can be estimated from the

recorded gaze data in the same way as fixation probability. The set of saccades $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$ was defined by $\mathbf{s}_i = \begin{pmatrix} \theta_i \\ r_i \end{pmatrix}$, where θ is the angle and r the log distance between each pair of successive fixations. Kernel density estimates were calculated as in Equation 2, but now replacing fixations \mathbf{x}_i with saccades \mathbf{s}_i . The log of distance was used because, typically for eye movements in natural scenes, the distribution of saccade amplitudes had a strong positive skew: the log transformation reduced this skew and improved the kernel density fit as measured by cross-validation. Separate bandwidths were used for angle and distance components, with optimal values again selected by cross-validation (free-viewing: $h_\theta = 8.0^\circ$, $h_r = 0.32$; search: $h_\theta = 13.8^\circ$, $h_r = 0.29$). The resulting saccade density estimate for free-viewing data is plotted in Figure 1d.

A second Monte Carlo simulation was performed to examine how preferences for different saccade amplitudes and directions affect gaze statistics. Two thousand sequences of saccades were generated, this time by pseudorandom sampling from the estimated saccade density. The starting point of the first saccade in each sequence was set to the centre of the display, matching the experimental data; subsequently, the endpoint of each simulated saccade became the start point for the next (example in Figure 2c).

Saccades that would result in a fixation outside the image boundaries were rejected and a new sample generated. Relative saccade metrics were calculated and frequencies averaged across simulation repetitions to obtain the distribution in Figure 2d.

Conditional probability estimation

An extension to the kernel density method (Hyndman, Bashtannyk, & Grunwald, 1996; Rosenblatt, 1969) provides a means of estimating the first-order conditional probability density of fixations within an image – that is, the probability of making an eye movement to location \mathbf{X} within an image, given that the current fixation is at location \mathbf{Y} .

This estimate is given by:

$$\hat{p}(\mathbf{x}_t = \mathbf{X} | \mathbf{x}_{t-1} = \mathbf{Y}) = \frac{\sum_i K_{h_1}(\mathbf{X} - \mathbf{x}_i) K_{h_2}(\mathbf{Y} - \mathbf{x}_{i-1})}{\sum_i K_{h_2}(\mathbf{Y} - \mathbf{x}_{i-1})}, \quad (5)$$

where K is defined as in Equation 3. Bandwidths h_1 and h_2 are again chosen by cross-validation (Holmes, Gray, & Isbell, 2007):

$$\arg \max_{h_1, h_2} \sum_I \left[\sum_i \log \sum_{j \neq i} K_{h_1}(\mathbf{x}_i^I - \mathbf{x}_j^I) K_{h_2}(\mathbf{x}_{i-1}^I - \mathbf{x}_{j-1}^I) \right]. \quad (6)$$

Optimal bandwidths obtained for free-viewing data were $h_1 = 1.27^\circ$, $h_2 = 3.30^\circ$, and for search $h_1 = 1.47^\circ$, $h_2 = 4.98^\circ$.

The conditional density differs from the fixation density (Equation 2) in that the fixation density has a fixed value for each location in an image, while the value of the conditional density changes depending on the start point of the hypothetical eye movement. Thus, the conditional density estimate captures not just overall preferences for different locations within an image (due to behavioral priority), but also the way that preference changes with current eye position (as a consequence of oculomotor biases). Critically, the conditional density does not incorporate any influence of preceding eye movements on selection, i.e., effects of gaze history. The conditional density estimate obtained for an example image is illustrated in Figure 3b.

To examine how the combination of fixation and saccade preferences affect gaze statistics, we performed a final Monte Carlo simulation. Two thousand sequences of eye movements, each matched in length to the experimental data, were generated for each image based on pseudorandom sampling from the conditional probability distribution given by Equation 5.

Specifically, an initial starting point (\mathbf{Y}_1) for each sequence was set to the center of the display, matching the experimental data. Then an endpoint \mathbf{X}_1 was generated by pseudorandom sampling from the conditional distribution $\hat{p}(\mathbf{x}_t | \mathbf{x}_{t-1} = \mathbf{Y}_1)$ for the image. This endpoint then became the starting location for the next saccade (i.e., $\mathbf{Y}_i = \mathbf{X}_{i-1}$) and the sampling process repeated to generate a new endpoint \mathbf{X}_2 , and so on, generating a simulated scan path.

An exact sampling from the conditional density estimate determined by Equation 5 was computationally impractical, so for the purposes of generating random samples we approximated the density by bilinear interpolation between values calculated for every element in a finely-spaced uniform grid of possible starting (\mathbf{Y}) and endpoint (\mathbf{X}) locations (2.56×10^6 grid points in total, grid separation 0.9°). Because the density estimate is necessarily smooth on scales at or below the kernel bandwidth, bilinear interpolation between these points provided a very close approximation to actual values. Further decreasing the grid point separation had no effect on results.

Relative saccade metrics were calculated for the simulated scanpaths in the same way as for recorded data (Equation 1), and frequencies averaged across simulation repetitions. The resulting distributions were then directly compared with those obtained from real data (Figures 4 and 5).

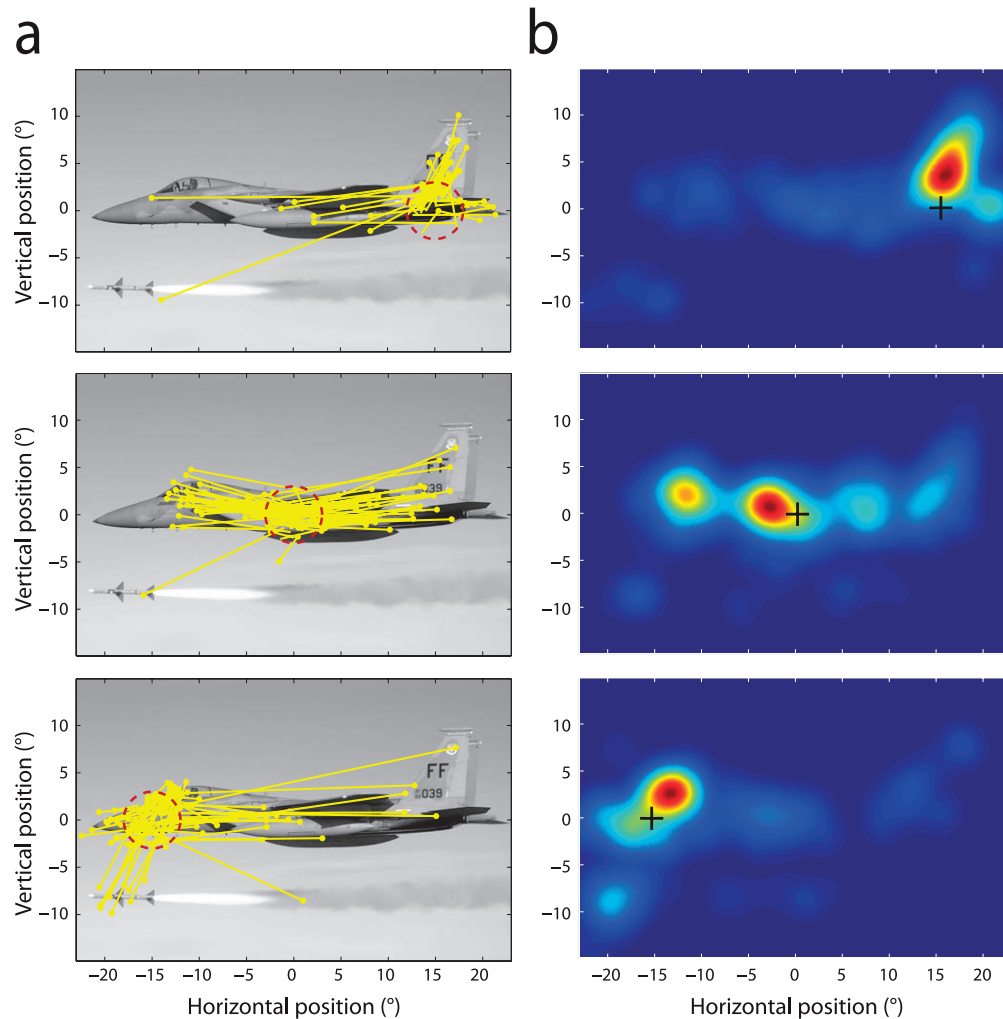


Figure 3. Conditional fixation probability captures instantaneous selection. **(a)** Examples of saccades (yellow) that immediately followed a fixation within one of three different areas of an example image (indicated by the red dashed circle in each of the three panels). Note that saccade selection depends on an interaction between *attentional priority* (directing gaze to salient locations in the image, such as the cockpit, missile or tail) and *oculomotor biases* (most evidently, a preference for smaller amplitude saccades). **(b)** Heat maps indicate the *conditional* probability of fixation within the same image for three different preceding fixation locations (denoted by black crosses). These conditional probabilities were estimated from all scanpaths recorded on the image.

Exploration efficiency

The “foraging facilitator” hypothesis (Klein, 2000) proposes that gaze history is incorporated into saccade selection so as to increase the efficiency with which visual scenes are examined. To compare the rate of exploration in real and simulated (memoryless) datasets, we examined how the proportion of total image area encompassed by a scanpath increased with number of fixations. Specifically, for each fixation made by a (real or simulated) subject on an image, we calculated the area of the smallest convex polygon (the *convex envelope*) containing all fixation points up to and including the present one (examples in Figure 8a).

To examine whether changes in exploration rate due to gaze history would have a behaviorally-meaningful

impact on the efficiency of search, we calculated the mean area of *target-present search* images explored by participants before identifying the target location. The number of fixations taken by real or simulated scanpaths to cover this threshold area in *target-absent search* and *free-viewing* images was taken as an estimate of how many fixations would have been required to find a target had it been present, and hence a measure of the efficiency of exploration.

Results

We recorded gaze position while participants viewed images of real-world scenes. An example image and

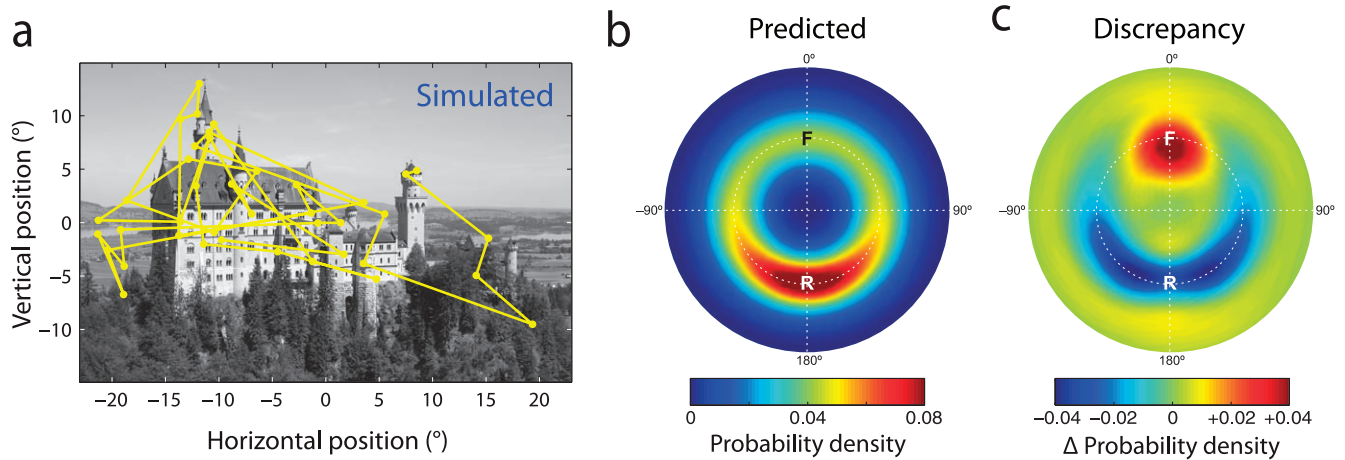


Figure 4. Gaze history influences saccade selection in free-viewing. **(a)** Typical scanpath generated by random sampling from the *conditional* fixation density (computed as in Figure 3b) estimated for another example image. The simulation takes into account *both* the distribution of attentional priority in the image and the oculomotor factors influencing saccade selection. **(b)** Relative saccade metrics in scanpaths simulated on the basis of conditional fixation density. This distribution corresponds to a prediction of the gaze statistics expected under *memoryless* saccade planning. Note prevalence of *return saccades* (R), despite simulation procedure having *no memory* for preceding eye movements. **(c)** Discrepancy between actual and predicted probability distributions, indicating the influence of gaze history on saccade planning. Hot colors correspond to facilitated eye movements, cold colors to inhibition. Note that return saccades are strongly inhibited compared to the prediction based on memoryless saccadic planning.

typical scanpath are shown in Figure 1a. The frequency of return eye movements was assessed by calculating the disparity in amplitude and direction between each successive pair of saccades in a sequence. Figure 1b displays the distribution of these *relative saccade metrics* in the scanpaths obtained from all participants and images.

Consistent with previous reports (Hooge et al., 2005; Smith and Henderson, 2009, 2011a), two peaks are observed in this distribution, corresponding to high frequencies of *forward* and *return* saccades. Forward saccades (labeled ‘F’ in the polar distribution shown in Figure 1b) are of approximately the same size and direction as the preceding eye movement; they made up 4.5% of total recorded eye movements ($|\Delta r| < 25\%$, $|\Delta\theta| < 30^\circ$). Return saccades (labeled ‘R’) are of approximately the same size but in the *opposite* direction to the preceding eye movement, i.e., they bring gaze back to the location of the previous fixation; they made up 4.3% of eye movements ($|\Delta r| < 25\%$, $|\Delta\theta| > 150^\circ$).

Both forward and return saccades were significantly more frequent than saccades of the same size in a neutral direction (90° to the previous eye movement, 1.9%; $t_{(19)} > 8.2$, $p < 0.001$). However, before we can infer that these classes of eye movement are actively facilitated in gaze planning, we must consider to what extent they may be generated by instantaneous processes, i.e., influences on selection that are independent of where the eyes have been before.

Attentional priority

In natural viewing, fixations are not randomly scattered, but instead preferentially directed towards certain regions of interest in a scene (Buswell, 1935; Henderson, Weeks, & Hollingsworth, 1999; Loftus & Mackworth, 1978; Mannan, Ruddock, & Wooding, 1995; Yarbus, 1967). To quantify the distribution of attentional priority in our images, we used kernel density estimation to obtain the *fixation density* in each scene: the underlying probability of a fixation falling at any given location, based on the total sample of recorded gaze positions.

As commonly observed, certain regions of an image were far more likely to be fixated than others (example in Figure 1c). To assess the extent to which this uneven distribution of attentional priority within an image could, by itself, be a cause of refixations, we generated a new set of *simulated* scanpaths by sampling gaze positions at random from the estimated fixation density (example in Figure 2a; see Methods).

The resulting simulated fixations are distributed between different regions of an image with the same probability observed in actual gaze, but critically each fixation is selected *independently* of the ones before, precluding any influence of gaze history. Nonetheless, the distribution of relative saccade metrics in the simulated dataset (Figure 2b) revealed a very high frequency of return saccades (15.3% of eye movements) compared to those of similar size in neutral directions (2.6%), which were in turn more frequent than forward saccades (1.5%).

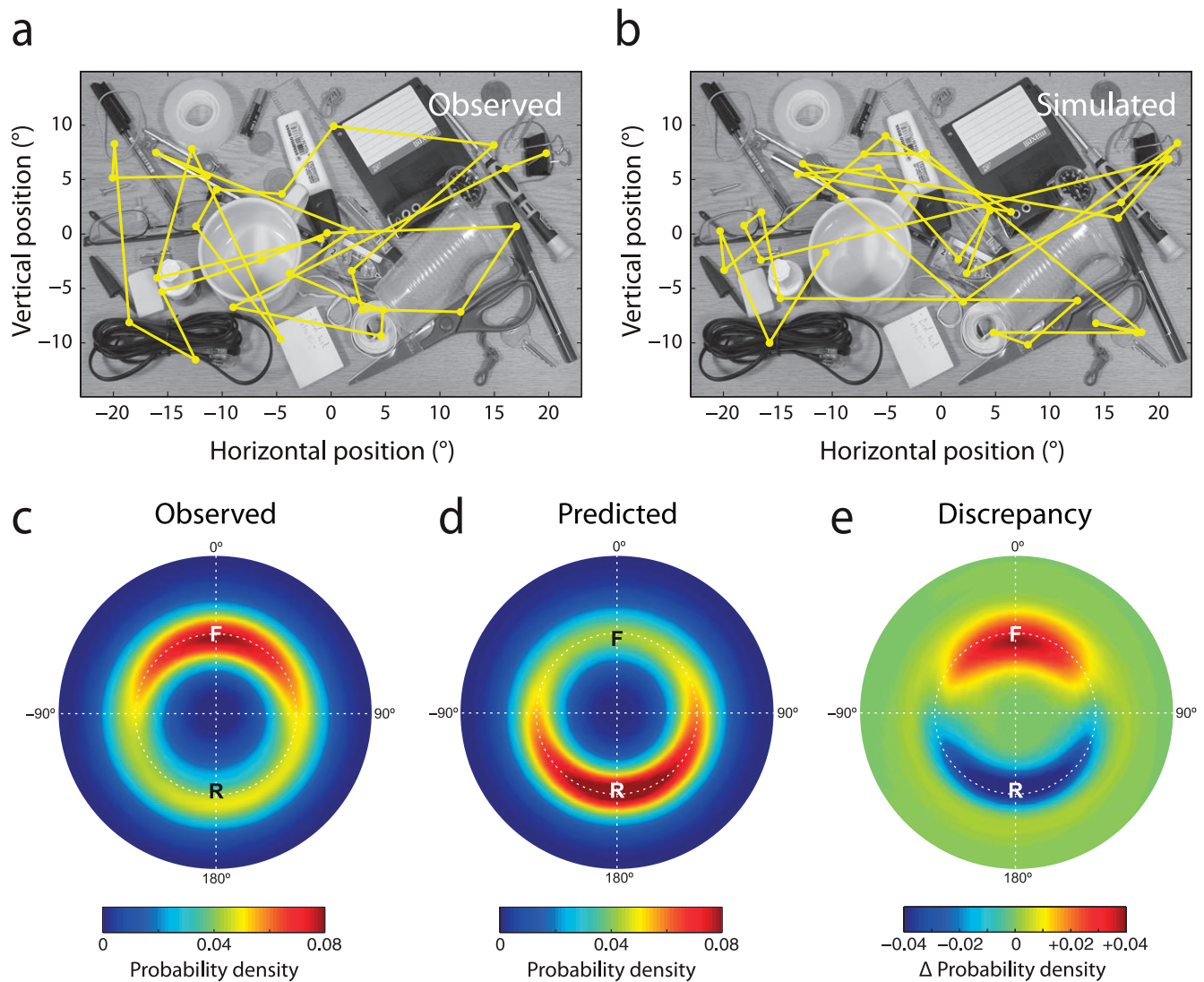


Figure 5. Inhibition of return in an active search task. **(a)** Example image and scanpath from the search task. Eye movements were recorded while participants searched for a specified item (a black drawing-pin/thumbtack, in this example). **(b)** Typical scanpath generated by random sampling from the conditional fixation density estimated for the same search image. **(c)** Heat map showing the distribution of *relative saccade metrics* observed during active search. Forward saccades are indicated by 'F' and return saccades by 'R'. **(d)** Relative saccade metrics in scanpaths simulated on the basis of conditional fixation density in search images. This distribution corresponds to a prediction of the gaze statistics expected under *memoryless* saccade planning. **(e)** Discrepancy between actual (c) and predicted (d) probability distributions, indicating the influence of gaze history on saccade planning during active search. Hot colors correspond to facilitated eye movements, cold colors to inhibition. Note that return saccades are strongly inhibited compared to the prediction based on memoryless saccadic planning.

This result demonstrates why comparing the frequency of return saccades with saccades to control locations matched for eccentricity does not provide an effective test for influences of gaze history (Hooge et al., 2005; Smith & Henderson, 2009). Biases favoring selection of certain locations within an image will tend to increase refixation frequency compared to control locations even if, as in our simulated scanpaths, each new fixation is chosen independently of those that went before.

Oculomotor bias

Despite visiting the same regions in an image with the same frequencies, the scanpaths simulated on the basis of fixation density alone (Figure 2a) do not resemble real gaze patterns. This is primarily because real scanpaths also display biases with respect to the amplitude and direction of saccades (Bahill, Adler, & Stark, 1975; Brandt, 1945; Foulsham, Kingstone, & Underwood, 2008; Gilchrist & Harvey, 2006; Tatler & Vincent, 2009). Consistent with previous reports,

certain saccade metrics were more frequent than others in our recorded data (Figure 1d).

Even though each image spanned more than 45° of visual angle, the data was dominated by saccades of small amplitude (median 4.3°). In addition, horizontal saccades were more than twice as frequent as vertical (51% of eye movements were within 30° of horizontal; 24% within 30° of vertical).

The origin of oculomotor biases in natural vision is debated, and could reflect low-level biomechanical factors, energetic considerations, limits on peripheral acuity or perceptual factors such as crowding (see Tatler & Vincent, 2009, for a detailed discussion). For the present purposes, biases in selecting saccade metrics are important in that they may make symmetrically-opposing eye movements more frequent than otherwise expected, and so contribute to the prevalence of return saccades. To test this hypothesis, we simulated a second set of scanpaths, this time by sampling eye movements at random from the observed distribution of saccade parameters (example in Figure 2c). The only constraint on saccade endpoints was that they remained within the boundaries of each image: the content of the scenes was ignored.

Even though each saccade was selected independently of those that had gone before, the resulting scanpaths again contained a higher proportion of return saccades (3.3%; Figure 2d), than those in neutral directions (2.5%). Forward saccades (2.7%) were also enhanced relative to neutral, though to a lesser extent.

This result demonstrates why methods that take into account the distribution of fixations in an image may still be insufficient to assess true influences of gaze history. For example, Smith and Henderson (2011b) compared observed refixation frequency to a baseline frequency obtained by shuffling the order of fixations. They found that a location was more likely to be fixated when it was the most recently visited location than at other times during the viewing period, and concluded that return saccades were actively facilitated. However, this same result was found in the present simulated data, even though each eye movement was selected independently of both image content and gaze history (frequency of returning to within 1° of preceding fixation: 3.3% unshuffled vs. 2.6% shuffled; $t_{(399)} = 6.0$, $p < 0.001$). Biases favoring certain metrics of eye movement over others are sufficient in themselves to generate a surplus of immediate refixations.

Memoryless saccade planning

The results of these simulations demonstrate that both attentional and oculomotor biases can act to increase the frequency of return saccades, even if the system of selection has no knowledge of preceding eye

movements. Figure 3a illustrates how these influences interact in the selection of gaze targets in natural vision. Each panel shows saccades initiated from a different region of one of our experimental images: the pattern of endpoints clearly follows the distribution of salient locations within the image, indicating the presence of attentional biases in selection. However, the distribution of endpoints is also very dependent on the starting location of the eye movement (differing from panel to panel), consistent with the influence of oculomotor factors, in particular a bias against large amplitude saccades. To obtain a baseline measure of return frequency against which the observed frequency can be compared, therefore, it is necessary to capture the combined effect of both fixation and saccade preferences on gaze planning. This result can be achieved by estimating the *first-order conditional probability density* of fixations in each image, i.e., the probability of choosing a saccade target location given the current fixation position (see Methods).

This conditional estimate, illustrated in Figure 3b, implicitly captures observers' preferences both for different saccade metrics and for different locations in the image, but crucially it does not incorporate any effect of preceding eye movements or fixations. If saccade planning is similarly *memoryless* (i.e., a Markov process), a simulated sequence of eye movements generated using the conditional estimate should be indistinguishable from a real saccade sequence. Any discrepancy between simulated and observed sequences will indicate an influence of memory on saccade planning.

Figure 4a shows an example from a set of scanpaths simulated by random sampling from the conditional fixation density, and Figure 4b the resulting distribution of relative saccade metrics. This distribution contains a single peak indicating a high predicted frequency of return saccades (6.5%, compared to 2.8% forward and 3.1% neutral).

The discrepancy between these simulated frequencies and those observed in actual gaze reveals the influence of gaze history on saccade selection (Figure 4c). Forward saccades were 63% more frequent than predicted for memoryless selection (SE 9%; $t_{(19)} = 7.1$; $p < 0.001$), indicating that they are actively facilitated. In contrast, the observed frequency of return saccades is 35% less than predicted (SE 4%; $t_{(19)} = 8.6$, $p < 0.001$), indicating that refixation is actively inhibited during free-viewing.

Saccade selection during active search

To test the generality of our findings, we repeated our analysis but now on gaze data recorded during a naturalistic search task, in which participants were

asked to search for everyday objects within a series of images of a cluttered desk (example image and scanpath in Figure 5a). In comparison to gaze patterns in free-viewing (where participants were asked simply to look at the pictures), saccades during search were on average of greater amplitude (median 7.5°). The bias towards horizontal saccades was similar to that in free-viewing (49% horizontal; 25% vertical).

Figure 5c shows the distribution of relative saccade metrics observed during active search. Unlike in free-viewing (Figure 1b), there was no peak in the distribution corresponding to return saccades; however, neither were these eye movements particularly infrequent (2.5%). To test for an influence of gaze history on selection, we again simulated scanpaths based on the conditional fixation density estimated for each search image (example scanpath in Figure 5b). As in free-viewing, scanpaths predicted on the basis of memoryless selection (Figure 5d) contained a much higher frequency of return saccades (7.0%).

The discrepancy between predicted and observed distributions (Figure 5a) confirmed that, as in free-viewing, forward saccades were systematically facilitated (85% more frequent than in memoryless selection, *SE* 12%; $t_{(19)} = 7.2$, $p < 0.001$) and return saccades inhibited (63% less frequent, *SE* 4%; $t_{(19)} = 16.6$, $p < 0.001$).

Amplitude specificity

To determine the frequencies of return and forward saccades in our data, we have used an operational definition that depends on the *relative* amplitude of each successive pair of saccades. Thus, symmetrically-opposing eye movements that return gaze to the previously fixated location are treated identically in our analysis irrespective of their amplitude. It is conceivable, however, that the processes governing saccadic selection may be influenced by the amplitude of the preceding eye movement. In particular, if the inhibition of return observed here were limited in its effect to small amplitude eye movements, then this limitation could bring into question its efficacy as a mechanism aiding scene exploration, which, it might be argued, depends more strongly on large amplitude saccades moving gaze between, rather than within, regions of a scene.

To examine whether the effects of gaze history were specific to particular amplitudes of saccade, we separately examined the frequency of forward and return eye movements following small (below median amplitude) and large (above median amplitude) saccades, in actual and simulated datasets. Consistent with the main results, analysis of the distribution of eye movements following small saccades revealed a signif-

icant facilitation of forward saccades (free-viewing: 59% more frequent than in memoryless selection; search: 81% more frequent; $t_{(19)} = 5.8$, $p < 0.001$) and a significant inhibition of return (free-viewing: 15% less frequent than in memoryless selection; search: 54% less frequent; $t_{(19)} = 6.8$, $p < 0.001$). Importantly, this same pattern was observed for the distribution of eye movements following large saccades (facilitation of forward saccades in free-viewing: 72%, in search: 95%; $t_{(19)} = 7.4$, $p < 0.001$; inhibition of return in free-viewing: 48%, in search: 68%; $t_{(19)} = 20.3$, $p < 0.001$). Thus we find no evidence to suggest that the effects of gaze history observed here are specifically associated with particular amplitudes of eye movement.

Comparison of spatial and temporal IOR

In previous studies, saccadic IOR has primarily been inferred from observations of saccade *latency* rather than frequency. To investigate the effect of gaze history on latency in our data, we calculated the duration of the fixation period between each successive pair of saccades.

Figure 6a and b show, for successive saccades matched in amplitude ($|\Delta r| < 25\%$), how median latency varied with relative direction of the eye movement. Gaze patterns recorded during free-viewing (Figure 6a) displayed a broadly-tuned *temporal* inhibition with its peak in the return direction (labeled ‘R’). Fixations preceding these return saccades ($|\Delta\theta| > 150^\circ$) lasted on average 300 ms. They were significantly prolonged relative to median fixation duration (blue dashed line; 247 ms; $t_{(19)} = 9.8$; $p < 0.001$), as well as compared to fixations preceding forward saccades ($|\Delta\theta| < 30^\circ$, 236 ms; $t_{(19)} = 10.0$; $p < 0.001$).

In comparison to free-viewing, fixations in active search were of significantly shorter duration overall (median 200 ms; $t_{(19)} = 12.8$, $p < 0.001$). A significant temporal IOR was again observed (Figure 6b). However, the effect was weaker than in free-viewing (return: 221 ms vs. forward: 194 ms; $t_{(19)} = 2.5$; $p = 0.022$).

For comparison, the effects of relative saccade direction on saccadic *frequency* (estimated from the conditional density as above) are shown in Figures 6c and d. The tuning of this spatial IOR was broadly similar to the temporal effect, with the strongest inhibition again centered on the return direction (labeled ‘R’) in both free-viewing and search tasks. In contrast to the latency effect, however, spatial IOR was stronger in active search than in free-viewing ($t_{(19)} = 4.8$, $p < 0.001$).

Figure 7 shows, for saccades in the return direction ($|\Delta\theta| > 150^\circ$), how latency (Figure 7a & b) and frequency (Figure 7c and d) varied with amplitude

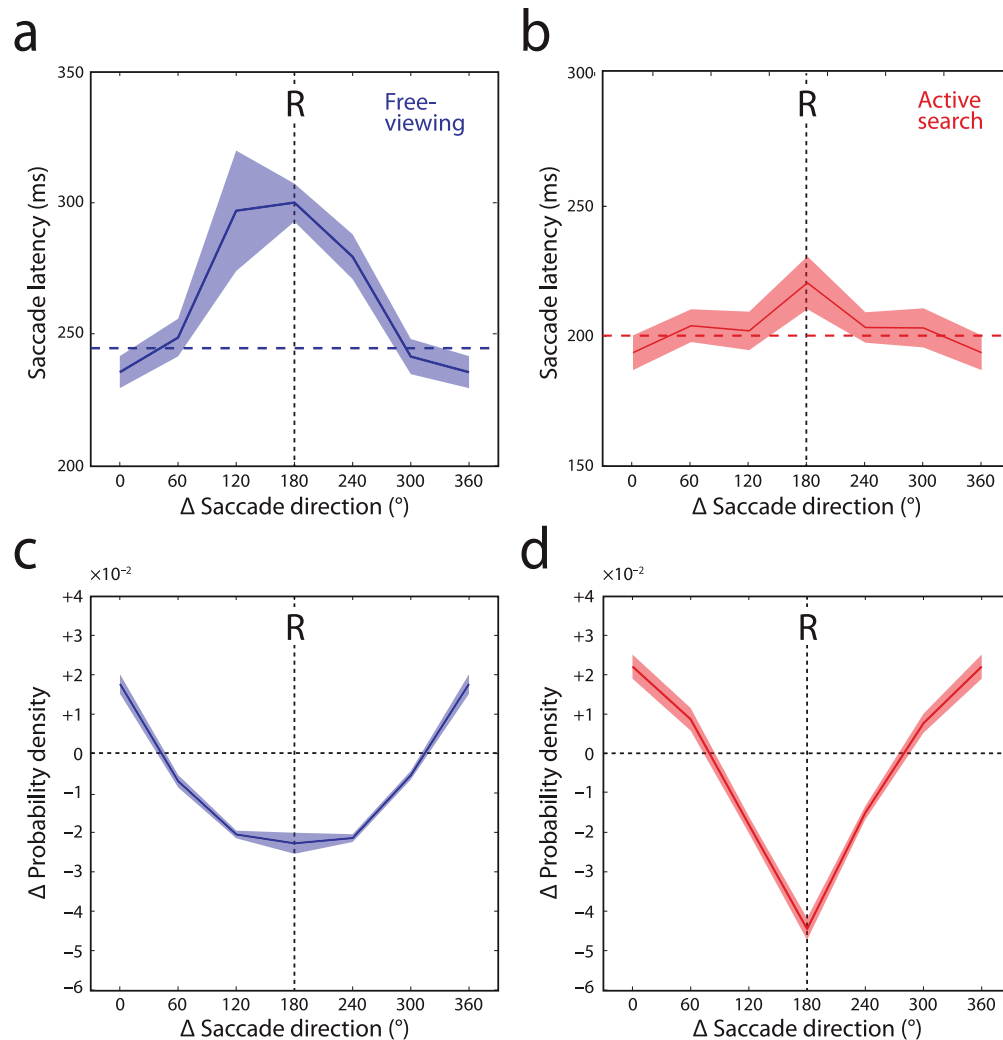


Figure 6. Directional specificity of temporal and spatial IOR. **(a, b)** Saccade latency as a function of direction relative to preceding eye movement ($\Delta\theta$), for saccades matched in amplitude ($|\Delta r| < 25\%$), recorded in free-viewing (a) and search tasks (b). Shaded region indicates ± 1 SE. Dashed horizontal lines indicate median latency of all saccades. ‘R’ indicates saccades in return direction (180°), i.e., those resulting in refixations. Note that fixation durations are prolonged prior to return saccades (*temporal IOR*). **(c, d)** Discrepancy between actual and simulated (*memoryless*) saccade frequencies as a function of direction relative to preceding eye movement ($\Delta\theta$), for saccades matched in amplitude ($|\Delta r| < 25\%$), in free-viewing (c) and in search (d). ‘R’ indicates refixations. Dotted horizontal line indicates equality. Negative values indicate that saccades are inhibited relative to memoryless selection (*spatial IOR*).

relative to the preceding eye movement. In free-viewing, the temporal IOR effect was strongest for return saccades matched in amplitude to the preceding eye movement (i.e., refixations $|\Delta r| < 25\%$), and declined with increasing disparity in size (Figure 7a).

The spatial IOR effect on saccade frequency (Figure 7c) was also strongest for accurate refixations, while saccades in the return direction but with very different amplitudes occurred *more* frequently than expected by chance (e.g., $|\Delta r| > 75\%$, 79% more frequent than in memoryless simulation; $t_{(19)} = 7.9$; $p < 0.001$).

In the active search task, the spatial IOR effect was similarly tuned to maximally inhibit refixations (Figure 7d), but a comparable tuning of the temporal IOR effect was not observed (Figure 7b). We found no

evidence for correlation at the participant level between the strength of temporal IOR (i.e., increase in return saccade latency) and the strength of spatial IOR (i.e., decrease in return saccade frequency) in either task ($r_s^2 < 0.13$, $p > 0.12$).

Consequences for visual exploration

The “foraging facilitator” hypothesis (Klein & MacInnes, 1999) proposes that return saccades are inhibited in order to facilitate rapid exploration of the visual environment. To investigate whether the effects of gaze history on saccade selection observed in the present study are sufficient to meaningfully influence

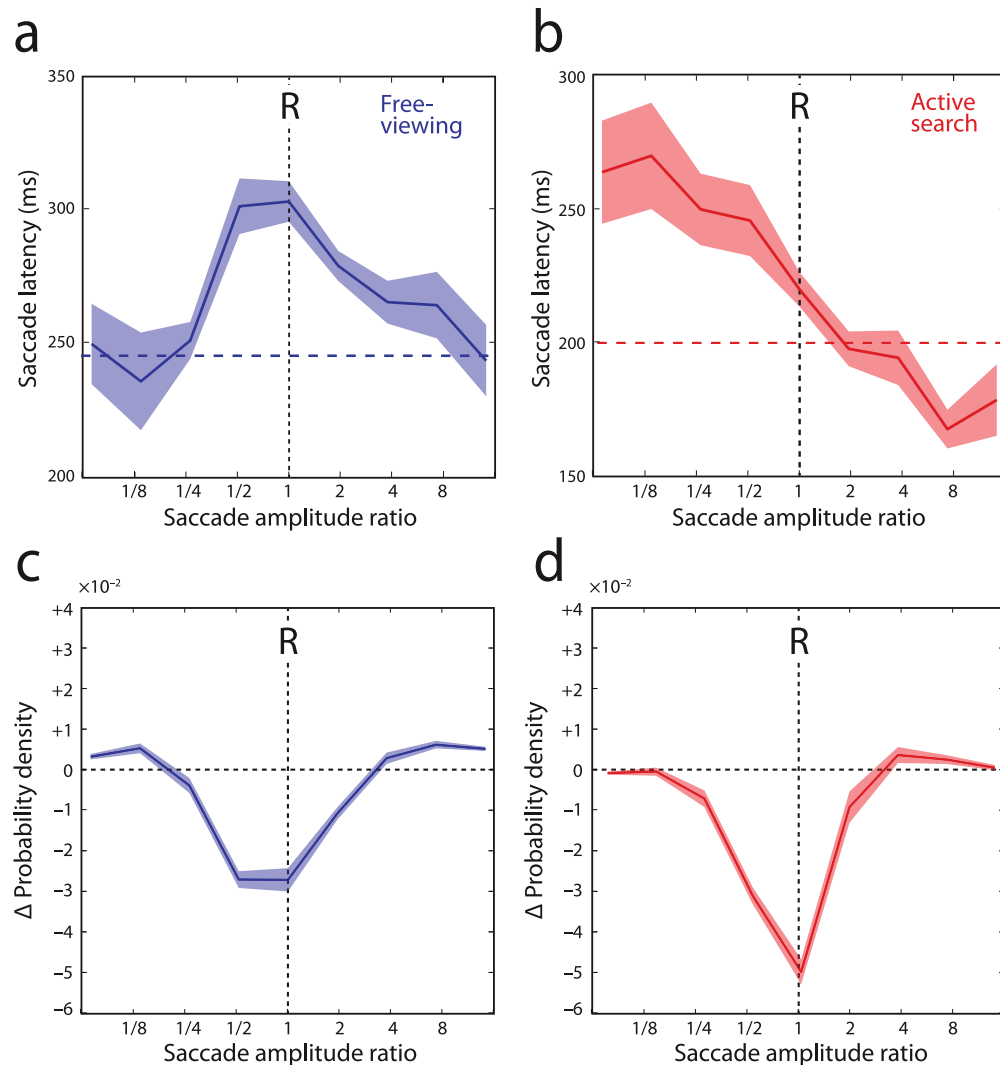


Figure 7. Amplitude specificity of temporal and spatial IOR. **(a & b)** Saccade latency as a function of amplitude relative to preceding eye movement (Δr), for saccades in the return direction ($|\Delta\theta| > 150^\circ$), recorded in free-viewing (a) and search tasks (b). Shaded region indicates ± 1 SE. Dashed horizontal lines indicate median latency of all saccades. ‘R’ indicates saccades matched in amplitude, i.e., those resulting in refixations. **(c & d)** Discrepancy between actual and simulated (*memoryless*) saccade frequencies as a function of amplitude relative to preceding eye movement (Δr), for saccades in the return direction ($|\Delta\theta| > 150^\circ$), in free-viewing (c), and in search (d). ‘R’ indicates refixations. Dotted horizontal line indicates equality. Negative values indicate that saccades are inhibited relative to memoryless selection.

exploratory behavior, we compared the rate of exploration in the recorded scanpaths with that predicted under *memoryless* selection.

Exploration rate was assessed based on the proportion of total image area enclosed by the scanpath (examples in Figure 8a). On search trials where the target was present, participants explored on average 24% of the total image area before identifying its location. We used this value as a threshold for assessing the behavioral significance of our results: the number of fixations required to explore this threshold area provides an estimate of how rapidly a search target

could be located and hence the efficiency of exploration.

Figure 8b displays the proportion of an image explored as a function of the number of fixations made on the image. Exploration was substantially faster in the search task (red solid line) than in free-viewing (blue solid line). Participants took on average 6.6 fixations ($SE 0.2$) to explore the target threshold area in the search task, compared to 18.2 fixations ($SE 0.9$) in free-viewing ($t_{(19)} = 13.6$; $p < 0.001$).

Dashed lines in Figure 8b show the area encompassed by simulated scanpaths, generated on the basis of the conditional fixation probability, i.e., excluding

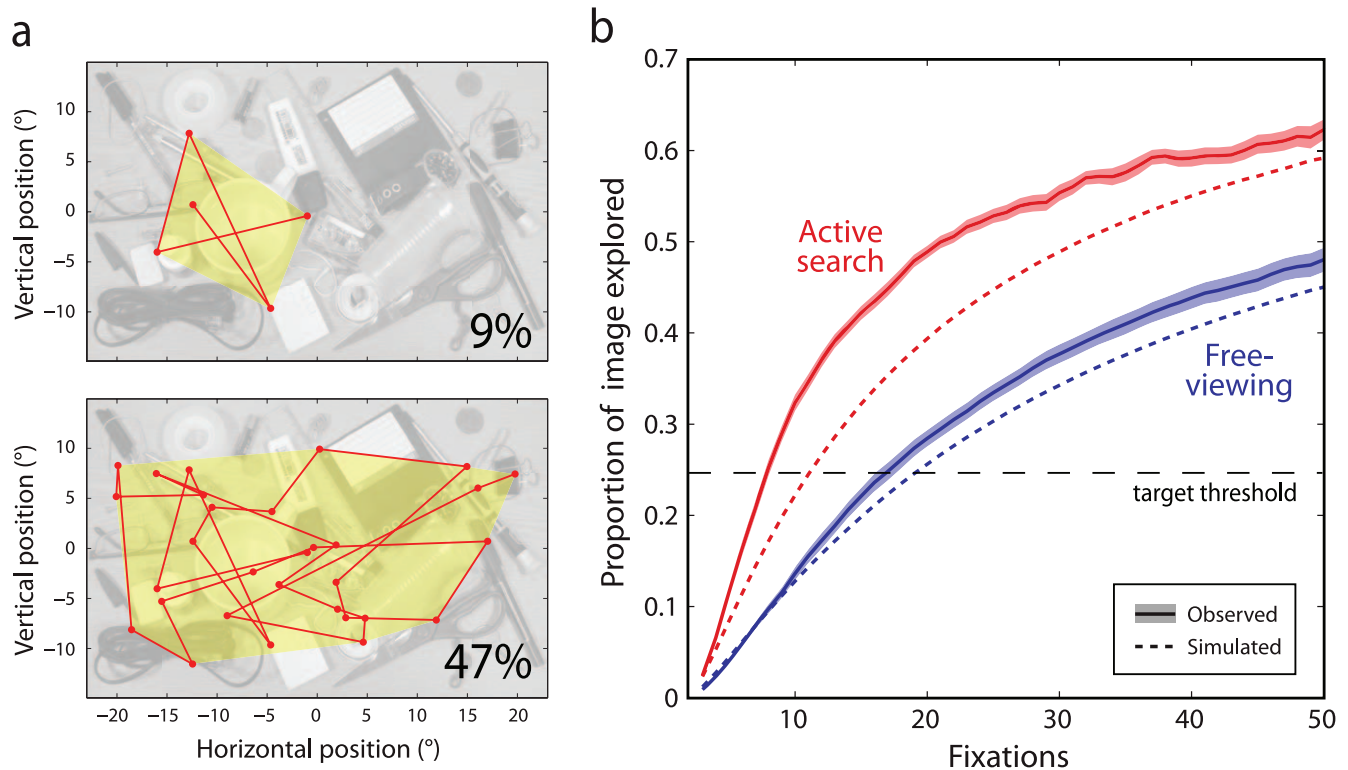


Figure 8. Memory for gaze history supports exploration in natural vision. **(a)** Proportion of image area explored at early (top) and late stages (bottom) of search, for a typical scanpath (red). In this example, the area enclosed by the scanpath (yellow shaded area) increases from 9% of total image area after the first five fixations (top) to 47% after thirty fixations (bottom). **(b)** Proportion of image area explored as a function of number of fixations, in recorded data (solid lines; shaded region indicates ± 1 SE) and *memoryless* simulation (dotted lines). Results for the free-viewing task are shown in blue, search task in red. Horizontal dashed line indicates mean area explored prior to locating target in search task.

influences of gaze history. Exploration in these simulated fixation sequences proceeded more slowly than observed in actual gaze, for both free-viewing and search.

The effect in free-viewing was modest, though statistically significant ($t_{(19)} = 2.2$, $p = 0.044$; 20.0 fixations to target threshold, SE 0.05). A strong effect was observed in the active search task (10.6 fixations, SE 0.05; $t_{(19)} = 17.4$, $p < 0.001$), where the memoryless simulation required 63% more fixations to explore the *target threshold area* than was observed for actual gaze.

Discussion

A controversial issue in eye movement research is the extent to which gaze planning is influenced by memory for previous fixations. It has been widely demonstrated that shifts of attention and gaze are initiated *more slowly* when directed back to a previously-examined location than to a novel one (Hooze & Frens, 2000; Klein, 1988; Spence & Driver, 1998; Tipper, Weaver, Jerreat, & Burak, 1994; Vaughan, 1984). A number of

authors (Klein, 2000; Klein & MacInnes, 1999; Posner and Cohen, 1984; Wang & Klein, 2010) have proposed that this temporal effect may be the observable consequence of a spatial inhibition of return that makes refixations *less likely*, and so supports exploration of new locations.

In this study, we have examined the process by which gaze locations are selected in natural vision. We considered eye movements recorded during both free-viewing of naturalistic images and active search among everyday objects. Two classes of regularity were observed: first, that certain regions of each image are fixated more often than others (Buswell, 1935; Henderson et al., 1999; Loftus & Mackworth, 1978; Mannan et al., 1995; Yarbush, 1967); and second, that certain amplitudes and directions of eye movement occur more frequently than others (Bahill et al., 1975; Brandt, 1945; Foulsham et al., 2008; Gilchrist & Harvey, 2006; Tatler & Vincent, 2009).

Previous studies of viewing and search of naturalistic scenes have concluded that spatial IOR does not play a significant role in natural vision (Hooze et al., 2005; Smith & Henderson, 2009, 2011a, 2011b). As in the current study (Figure 1b), these authors observed that

return saccades (eye movements returning gaze to the location of the previous fixation) were *more frequent* than saccades in neutral directions.

However, for IOR to facilitate visual exploration, it is not necessary for refixations to be entirely prevented, only made *less likely*. In the current study, we tested this hypothesis by estimating the frequency of return saccades expected if saccade planning were memoryless. We found that the instantaneous influences on gaze planning described above were sufficient to generate substantial numbers of return saccades, even though the selection process had no knowledge of gaze history. This result comes about for two reasons.

First, the location of a previous fixation will in general have above-average priority (salience and/or relevance) within the scene: in the absence of memory-based inhibition, an observer is therefore likely to *refixate* a region for the same reason they first fixated it (Figures 2a and b). Second, observers' systematic bias towards selecting saccades of particular amplitudes will make symmetrically-opposing eye movements (i.e., return saccades) more frequent than otherwise expected (Figures 2c and d).

Comparison with observed frequencies revealed that return saccades were considerably less frequent than predicted for a memoryless process (Figures 4 and 5), indicating that, contrary to previous reports, IOR significantly influences gaze planning in natural vision. This comparison also revealed that the high frequency of forward saccades (i.e., eye movements of similar amplitude and direction to the one before) observed in naturalistic viewing cannot be explained by instantaneous processes, and therefore represents a further influence of memory on saccade planning.

The inhibitory effect was broadly tuned to both the relative amplitude and direction of successive saccades, such that the most strongly inhibited eye movements were those that would generate a precise refixation at the previously-visited location (Figures 6 and 7). Consistent with the proposed role for IOR, as a “foraging facilitator,” inhibition was stronger and more precisely tuned in the active search task, when behavioral goals necessitated rapid exploration, than in the free-viewing task.

In previous research, prolonged fixations prior to return eye movements have been interpreted as indirect evidence that these saccade plans are inhibited. In the present study, we were able to make direct comparisons between this *temporal* IOR effect and the *spatial* IOR expressed in the frequency of return saccades. In free-viewing, analysis of saccade latencies revealed a strong temporal IOR with a broad tuning that mirrored the changes observed in saccade frequency. However, the enhancement of spatial IOR observed in active search was not paralleled in saccade timing: fixation durations

were on average shorter during search and only weakly influenced by the preceding eye movement.

The functional significance of IOR depends on its ability to facilitate rapid exploration of the visual environment, which in turn depends on the ability to prevent refixations, not delay them. Thus, temporal IOR may be considered a secondary consequence of an inhibitory process primarily aimed at reducing the *frequency* with which return saccades are generated. As such, the temporal effect may be distorted or obscured by other, stronger influences on saccadic latency, such as motivation or expected reward (Reddi & Carpenter, 2000; Takikawa, Kawagoe, Itoh, Nakahara, & Hikosaka, 2002). This may explain why, despite similarities at the group level, individual differences in the strength of temporal IOR were not correlated with differences in spatial IOR in this study.

To act as a foraging facilitator, the effect of previous saccades on gaze planning must be strong enough to meaningfully affect the efficiency of visual exploration. In our ecologically motivated search task, participants scanned a cluttered desk for everyday objects: finding a target item required on average visual exploration of one quarter of the total desk surface. Our simulation of saccade planning based on instantaneous influences alone indicated that memory for gaze history was responsible for an increase in the rate of exploration, corresponding to a behaviorally-meaningful (~40%) reduction in the number of fixations required to achieve this goal (Figure 8).

The enhanced spatial IOR observed when participants were engaged in search compared to free-viewing suggests that the strength of return inhibition may be to some extent under top-down control. This hypothesis is also supported by a previous study (Farrell et al., 2010) which observed modulation of return latencies by knowledge of task statistics. In this previous study, subjects made short sequences of eye movements instructed by cues presented at fixation, each cue indicating the direction of the next saccade. When cued directions were random, return saccades were initiated more slowly than those in new directions; however, when cued directions consistently favored the return direction over a block of trials, this temporal IOR was abolished.

Importantly, fitting a rise-to-threshold model to saccade latencies in this task suggested two distinct components were involved in generating the temporal IOR effect: a slower accumulation for return versus neutral directions, independent of task statistics, and a lower threshold for selecting high-probability versus low-probability targets. This result illustrates the more general point that return inhibition may not be a monolithic process, but may instead arise from the combined contribution of multiple mechanisms.

At the most basic level, return inhibition could arise from a simple trajectory bias in saccade generation,

favoring eye movements that continue in the direction of the last saccade over those that reverse direction. The high frequency and low latency of forward saccades observed here and in previous studies suggests that a forward bias is an important component of saccade planning in natural vision. Such a tendency has been referred to as “saccadic momentum” (Smith & Henderson, 2009), although it is important to emphasize that this does not refer to a physical property of the oculomotor apparatus. A neural representation of the direction of the preceding eye movement must be maintained in some form, although this memory might be low-level, for example encoded in residual activity in superior colliculus (Wang, Satel, Trappenberg, & Klein, 2011).

Smith and Henderson (2009) found evidence for a trajectory bias in the latencies of saccades in natural scenes: fixations were prolonged prior to eye movements in the reverse direction, even when the saccade amplitude was too long or short to produce an exact refixation. However, they also found that this directional bias alone could not account for the specificity with which temporal IOR targets refixations: the longest latencies were observed for saccades matched both in amplitude and direction to the return location. While the prevalence of return saccades in their data led these authors to conclude that IOR influenced their latency but not their frequency, the present results demonstrate a spatial inhibition that is similarly tuned to both the amplitude and direction of the preceding eye movement, maximally inhibiting the selection of accurate refixations (Figures 6 & 7).

In contrast to research using naturalistic images, studies of oculomotor search in sparse artificial displays (Boot, McCarley, Kramer, & Peterson, 2004; Gilchrist & Harvey, 2000; Keech & Resca, 2010; McCarley, Wang, Kramer, Irwin, & Petersen, 2003; Motter & Holsapple, 2007; Peterson, Kramer, Wang, Irwin, & McCarley, 2001) have tended to emphasize the importance of higher-level representations in preventing return to previously-examined locations. McCarley et al. (2003) used a gaze-contingent display to present observers with just two choices of target for each eye movement. In this situation, gaze was preferentially directed to new locations over locations examined even several fixations previously. Furthermore, refixation was least likely for targets that had remained visible during the intervening fixations, supporting a role for persistent object representations of the kind typically associated with visual short-term memory, in addition to purely spatial inhibition.

A follow-up study (Boot et al., 2004) found that the tendency to select new saccade targets could be only partially overcome by removing the search element of the task and instructing subjects to intentionally saccade to old items. Like the Farrell et al. study, this result suggests that return inhibition has both automatic

components that operate irrespective of task demands, and intentional or strategic components that may be voluntarily suspended or reversed when refixation is required. Indeed, it is generally recognized that refixation serves a useful purpose in many real-world situations, by supporting reexamination of scene elements that were incompletely processed on a previous fixation, or (in dynamic environments) may have changed since last examined. It is possible such intentional refixations may have contributed to the gaze patterns recorded in the present study; however the dominant effect of previous fixations observed here is an inhibition of return compared to memoryless selection.

The present findings resolve the conflict between results based on sparse search arrays that have largely supported the existence of a spatial inhibition of return, and observations of oculomotor behavior in viewing naturalistic images that seemingly refute it (Hooge et al., 2005; Smith & Henderson, 2009, 2011a, 2011b). According to our results, the differing conclusions drawn by these previous studies should not be taken as indicating a qualitative difference in gaze behavior between artificial and natural environments. Rather, sparse search arrays and gaze-contingent displays allow the experimenter to directly control the salience and proximity of targets, whereas a more sophisticated analysis of the kind presented here is necessary to account for these factors in natural images. Once a baseline for these instantaneous influences is established, gaze patterns in natural images also display a strong inhibition of return saccades.

In common with previous studies, one of the two data sets used in our analysis consisted of gaze patterns recorded while freely viewing static photographs of real scenes. While important here for consistency with previous studies, the “picture-viewing” paradigm has been criticized by some authors as a relatively poor surrogate for vision in real environments (Henderson, 2003; Tatler, Hayhoe, Land, & Ballard, 2011). Many of these criticisms are mitigated by our second data set of eye movements in a naturalistic search task—in particular, unlike the free-viewing condition, participants were given a clearly-defined goal of relevance to everyday life (locating a specified object on a cluttered desk), and all the images were presented at a realistic scale and distance from the observer. However, this condition could still be criticized as “unnatural” on several counts, including the vertical orientation of the image, its sudden onset, absence of binocular depth cues, and the fact that subjects were required to keep their head and body still and could not physically interact with the objects. Inevitably, no individual experiment can capture the full range of situations and goals the human visual system may have to contend with in the natural world. However, it seems reasonable to conclude from the present results that inhibition of

saccadic return is an important component of gaze planning in natural as well as artificial settings.

The present results have important implications for computational models based on a saliency or priority map (Bruce & Tsotsos, 2009; Itti & Koch, 2000, 2001; Kanan, Tong, Zhang, & Cottrell, 2009; Koch & Ullman, 1985; Parkhurst, Law, & Neiber, 2002; Torralba, Oliva, Castelhano, & Henderson, 2006; Wolfe, 1994). A number of saliency-based models already implement some form of spatial IOR, for example by transiently inhibiting a region of the salience map centered on the current focus of attention (Itti & Koch, 2000; Koch & Ullman, 1985). Indeed, local inhibition at the current focus of attention is essential for saliency-based models to shift attention at all in static scenes—without it attention would remain fixed at the point of maximum saliency in the image. By contrast, the role of IOR in discouraging *return* to a previously examined location has been given less consideration in these models.

The earliest saliency models appear to have implemented IOR as a complete inhibition of activity in the saliency map within a fixed region surrounding the focus of attention. In a model of saccade selection, this would have the effect of entirely preventing refixations, an outcome inconsistent with empirical gaze patterns. While subsequent iterations have introduced more graded inhibition, the intensity and spatial extent of IOR applied in these models does not have a strong empirical basis. The present results have the potential to inform computational models of gaze planning by providing quantitative estimates of the strength and specificity of return inhibition in natural viewing. Additionally, to our knowledge no current saliency-based model of gaze selection reproduces the high frequency of forward saccades observed here. Incorporating this forward bias will be essential to capturing the process of gaze control in natural vision.

Acknowledgments

We thank Petroc Sumner for comments on the manuscript, Martin Bays for helpful discussion, and Louise Marshall for additional data collection. This work was supported by the Wellcome Trust and the NIHR CBRC at UCL/UCLH.

Commercial relationships: none.

Corresponding author: Paul M. Bays

Email: p.bays@ucl.ac.uk.

Address: Sobell Department of Motor Neuroscience and Movement Disorders, UCL Institute of Neurology, Queen Square, London, UK.

References

- Bahill, A. T., Adler, D., & Stark, L. (1975). Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative Ophthalmology & Visual Science*, *14*(6):468–469, <http://www.iovs.org/content/14/6/468>. [Article]
- Bichot, N. P., & Schall, J. D. (2002). Priming in macaque frontal cortex during popout visual search: feature-based facilitation and location-based inhibition of return. *Journal of Neuroscience*, *22*, 4675.
- Bisley, J. W., & Goldberg, M. E. (2003). Neuronal Activity in the Lateral Intraparietal Area and Spatial Attention. *Science*, *299*, 81–86
- Boot, W. R., McCarley, J. S., Kramer, A. F., & Peterson, M. S. (2004). Automatic and intentional memory processes in visual search. *Psychonomic Bulletin & Review*, *11*, 854–861.
- Brandt, H. F. (1945). The psychology of seeing. *The Philosophical library*. Oxford, UK: Philosophical Library.
- Bruce, N. D. B., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, *9*(3):5, 1–24, <http://www.journalofvision.org/content/9/3/5>, doi: 10.1167/9.3.5. [PubMed] [Article]
- Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Dorris, M. C., Klein, R. M., Everling, S., & Munoz, D. P. (2002). Contribution of the primate superior colliculus to inhibition of return. *Journal of Cognitive Neuroscience*, *14*, 1256–1263.
- Farrell, S., Ludwig, C. J. H., Ellis, L. A., & Gilchrist, I. D. (2010). Influence of environmental statistics on inhibition of saccadic return. *Proceedings of the National Academy of Sciences*, *107*, 929–934.
- Fecteau, J. H., & Munoz, D. P. (2006). Saliency, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, *10*, 382–390.
- Fecteau, J. H., & Munoz, D. P. (2011). Correlates of capture of attention and inhibition of return across stages of visual processing. *Journal of Cognitive Neuroscience*, *17*, 1714–1727.
- Foulsham, T., Kingstone, A., & Underwood, G. (2008). Turning the world around: Patterns in saccade direction vary with picture orientation. *Vision Research*, *48*, 1777–1790.
- Gardiner, C. (2010). *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Berlin: Springer.

- Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, *10*, 1209–1212.
- Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scan paths in visual search. *Visual Cognition*, *14*, 704–715.
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, *391*, 481–484.
- Habbema, J. D. F., Hermans, J., & Van Den Broek, K. (1974). A stepwise discriminant analysis program using density estimation. *Compstat*, *1974*, 101–110.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504.
- Henderson, J. M., Weeks, P. A., & 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.
- Holmes, M. P., Gray, A. G., & Isbell, Jr. C. L. (2007). Fast nonparametric conditional density estimation in uncertainty in artificial intelligence. (UAI).
- Hooge, I. T., & Frens, M. A. (2000). Inhibition of saccade return (ISR): spatio-temporal properties of saccade programming. *Vision Research*, *40*, 3415–3426.
- Hooge, I. T. C., Over, E. A. B., van Wezel, R. J. A., & Frens, M. A. (2005). Inhibition of return is not a foraging facilitator in saccadic search and free viewing. *Vision Research*, *45*, 1901–1908.
- Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *394*, 575–577.
- Hyndman, R. J., Bashtannyk, D. M., & Grunwald, G. K. (1996). Estimating and Visualizing Conditional Densities. *Journal of Computational and Graphical Statistics*, *5*, 315–336.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–204.
- Kanan, C., Tong, M.H., Zhang, L., & Cottrell, G.W. (2009). SUN: Top-down saliency using natural statistics. *Visual Cognition*, *17*, 979–1003.
- Keech, T., & Resca, L. (2010). Eye movement trajectories in active visual search: Contributions of attention, memory, and scene boundaries to pattern formation. *Attention, Perception, & Psychophysics*, *72*, 114–141.
- Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature*, *334*, 430.
- Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*, 138–147.
- Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, *10*, 346–352.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.
- Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 565–572.
- Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision*, *9*, 363–386.
- McCarley, J. S., Wang, R. F., Kramer, A. F., Irwin, D. E., & Peterson, M. S. (2003). How much memory does oculomotor search have? *Psychological Science*, *14*, 422–426.
- McPeck, R. M., & Keller, E. L. (2002). Saccade target selection in the superior colliculus during a visual search task. *Journal of Neurophysiology*, *88*, 2019.
- Mirpour, K., Arcizet, F., Ong, W. S., & Bisley, J. W. (2009). Been there, seen that: A neural mechanism for performing efficient visual search. *Journal of Neurophysiology*, *102*, 3481–3491.
- Motter, B. C., & Holsapple, J. (2007). Saccades and covert shifts of attention during active visual search: Spatial distributions, memory, and items per fixation. *Vision Research*, *47*, 1261–1281.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123.
- Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science*, *12*, 287–292.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and performance X: Control of language processes*, *32*, 531–556.
- Reddi, B. A. J., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature Neuroscience*, *3*, 827–830.
- Reuter-Lorenz, P. A., Jha, A. P., & Rosenquist, J. N. (1996). What is inhibited in inhibition of return? *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 367–378.
- Rosenblatt, M. (1969). Conditional probability density and regression estimators. In P. R. Krishnaiah

- (Ed.), *Multivariate analysis 2* (pp. 25–31). Amsterdam: North-Holland.
- Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annual Reviews in Neuroscience*, *22*, 241–259.
- Silverman, B. W. (1986). Density estimation for statistics and data analysis. London: Chapman & Hall.
- Smith, T., & Henderson, J. (2009). Facilitation of return during scene viewing. *Visual Cognition*, *17* (6), 1083–1108.
- Smith T. J., & Henderson, J. M. (2011a). Looking back at Waldo: Oculomotor inhibition of return does not prevent return fixations. *Journal of Vision* *11*(1):3, 1–11, <http://www.journalofvision.org/content/11/1/3>, doi:10.1167/11.1.3. [PubMed] [Article]
- Smith, T. J., & Henderson, J. M. (2011b). Does oculomotor inhibition of return influence fixation probability during scene search? *Attention Perception and Psychophysics*, *73*, 2384–2398.
- Spence, C., & Driver, J. (1998). Auditory and audiovisual inhibition of return. *Perception & Psychophysics*, *60*, 125–139.
- Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., & Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Experimental Brain Research*, *142*, 284–291.
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, *11*(5):5, 1–23, <http://www.journalofvision.org/content/11/5/5>, doi:10.1167/11.5.5. [PubMed] [Article]
- Tatler, B. W., & Vincent, B. T. (2009). The prominence of behavioural biases in eye guidance. *Visual Cognition*, *17*, 1029–1054.
- Thompson, K. G., & Bichot, N. P. (2005). A visual salience map in the primate frontal eye field. *Progress in Brain Research*, *147*, 251–262.
- Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology-Human Perception and Performance*, *20*, 478–498.
- Torralba, A., Oliva, A., Castelhana, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*, 766–786.
- Vaughan, J. (1984). Saccades directed at previously attended locations in space. In *Theoretical and Applied Aspects of Eye Movement Research, Selected/Edited Proceedings of The Second European Conference on Eye Movements*. (pp. 143–150). Amsterdam: North-Holland.
- Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research*, *50*, 220–228.
- Wang, Z., Satel, J., Trappenberg, T. P., & Klein, R. M. (2011). Aftereffects of saccades explored in a dynamic neural field model of the superior colliculus. *Journal of Eye Movement Research*, *4*, 1–16.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of guided search. *Psychonomic Bulletin & Review*, *1*, 202–238.
- Yarbus, A. L. (1967). Eye movements and vision. New York: Plenum Press.