

**The prominence of behavioural biases in eye guidance**

Benjamin W. Tatler\* & Benjamin T. Vincent

University of Dundee, UK

RUNNING HEAD: Behavioural biases in eye guidance

KEYWORDS: behavioural bias, systematic tendency, saccade amplitude, saccade direction, salience, fixation

\* Corresponding author: [b.w.tatler@activevisionlab.com](mailto:b.w.tatler@activevisionlab.com) or [b.w.tatler@dundee.ac.uk](mailto:b.w.tatler@dundee.ac.uk)

Address for correspondence:

School of Psychology,

University of Dundee,

Dundee, DD1 4HN,

Scotland, UK

Tel: +44 (0) 1382 388260

Fax: +44 (0) 1382 229993

AUTHOR NOTE: Both authors contributed equally to this manuscript.

ABSTRACT

When attempting to understand where people look during scene perception, researchers typically focus on the relative contributions of low- and high-level cues. Computational models of the contribution of low-level features to fixation selection, with modifications to incorporate top-down sources of information have been abundant in recent research. However, we are still some way from a model that can explain many of the complexities of eye movement behaviour. Here we show that understanding biases in *how* we move the eyes can provide powerful new insights into the decision about where to look in complex scenes. A model based solely on these biases and therefore blind to current visual information outperformed popular salience-based approaches. Our data show that incorporating an understanding of oculomotor behavioural biases into models of eye guidance is likely to significantly improve our understanding of where we choose to fixate in natural scenes.

## INTRODUCTION

Successfully completing many forms of behaviour requires that humans look in the right place at the right time. Ballard and colleagues described this as a “do-it-where-I’m-looking” visual strategy for completing complex tasks (Ballard et al., 1992); a finding that has been replicated across a range of studies of natural behaviour (e.g., Hayhoe et al., 2003; Land & Hayhoe, 2001; Land et al., 1999; Pelz & Canosa, 2001).

One reason why we look at the location we are interested in gathering information from is that the human retina evolved such that high quality vision is restricted to the small ( $\sim 2^\circ$ ) fovea at the centre of vision. For many visually-guided behaviours the coarse information from peripheral vision is insufficient, thus requiring mechanisms to direct the foveae to appropriate locations. This has generated a large volume of research aimed at understanding how the eyes are guided (since Buswell, 1935).

### **Feature-based accounts of eye guidance**

One promising approach to understanding eye guidance has been to suppose that ‘basic’ visual features such as contrast, edges, colour and motion, are extracted from the visual scene and used to direct the eyes (e.g., Wolfe & Horowitz, 2004). Support for this notion can be found from visual search paradigms in psychophysics, where manipulating visual features has clear consequences on the deployment of visual attention (Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Wolfe, 1998). This notion of basic features guiding attention has been extended to natural scene viewing and formalized in models of eye guidance. One prominent such model is the visual salience model of Itti, Koch and colleagues (Itti & Koch, 2000; Itti, Koch & Niebur, 1998; Koch & Ullman, 1985). Here salience is operationalised as the output of a competitive process between a set of basic features (colour-, orientation- and luminance-contrast) in order to produce an overall salience map of the scene. Eye guidance then unfolds from this using a winner-takes-all selection of the most salient location in the scene, with transient inhibition of fixated locations to avoid the model becoming stuck.

Despite the prominence of feature-based accounts of eye guidance in recent years, empirical evaluations of such models have shown that these are disappointingly poor at accounting for human fixation selection (e.g., Henderson et al., 2007; Tatler, 2007; Tatler, Baddeley & Gilchrist, 2005; Tatler, Baddeley and Vincent, 2006). In particular, when the behavioural task is manipulated, feature-based models can fail almost completely (e.g., Einhäuser, Rutishauser & Koch, 2008a; Foulsham & Underwood, 2008; Underwood & Foulsham, 2006; Underwood, Foulsham, van Loon, Humphreys & Boyce, 2006).

Selectively weighting the different feature channels in Itti’s salience model (Navalpakkam & Itti, 2005) is one way to potentially improve feature-based accounts and incorporate some degree of high-level modulation (by effectively supplying top-down knowledge of the target of a visual search). However, even this modification of the salience model is very limited for finding real objects in images of natural scenes (Vincent, Troscianko & Gilchrist, 2007).

More success has been found by incorporating top-down knowledge of where targets are likely to be found in natural scenes (Ehinger, Hidalgo-Sotelo, Torralba & Oliva, this issue; Torralba, 2003; Torralba, Oliva, Castelhana & Henderson, 2006). Scene gist is used to categorise the scene and look up a spatial probability distribution of where the target is likely to be found. The spatial probability distribution is used to constrain the feature-level computations to these likely locations.

Despite continued research effort, we remain some way from a coherent understanding of the factors that underlie saccade target selection when viewing natural scenes and during natural behaviour.

### **Re-phrasing eye guidance probabilistically**

The fundamental question at the heart of any account of eye guidance must be to understand the moment-to-moment relocation of gaze: That is, where will the eyes select as the target of the next fixation? A convenient way to phrase this question is in the language of probabilities. In general we can say that we are interested in knowing the probability of making a saccade to a location based upon all the information that the oculomotor system has available to it:  $P(\text{saccade}|\text{data})$ . Calculating this spatial probability distribution (or map) directly is not necessarily impossible, but by using Bayes Theorem we can break this down into simpler components:

$$P(\text{saccade} | \text{data}) = \frac{P(\text{data} | \text{saccade})}{P(\text{data})} P(\text{saccade}) \quad (\text{eq. 1})$$

The beauty of this approach is that the data could come from a variety of sources such as simple feature cues, derivations such as Itti's definition of salience, object- or other high-level sources. While this approach is extremely general and flexible in that manner, for the present study and for comparability with the studies discussed above, we will consider saccade target selection on the basis of the lower-level cues of visual salience or simple visual features (in this case, edges).

The first right hand term in equation 1,  $P(\text{data}|\text{saccade})/P(\text{data})$  describes how the visual data might be involved in saccade target selection. Specifically,  $P(\text{data}|\text{saccade})$  is the likelihood of particular visual data (say, particular image features) occurring at a saccade target location, and  $P(\text{data})$  is the probability distribution of these visual data occurring in the environment. As such, dividing  $P(\text{data}|\text{saccade})$  by  $P(\text{data})$  effectively controls for the natural abundance of particular features within scenes. For example, if yellow items are commonly fixated then one may *initially* infer that yellow items predict fixations, but if yellow items are very common in the scene then yellow is a less effective predictor of eliciting fixations.

What is described in this first term bears a close resemblance to approaches previously employed to evaluate the possible involvement of visual features in eye guidance: Visual feature content at fixation – a measure of  $P(\text{data}|\text{saccade})$  – is compared to features at control locations – an approximation of  $P(\text{data})$  – and any differences are taken to imply non-random selection with respect to the visual feature under investigation (e.g., Mannan, Ruddock & Wooding, 1997; Reinagel & Zador,

1999; Krieger, Rentschler, Hauske, Schill & Zetzsche, 2000; Parkhurst et al., 2002; Parkhurst & Neibur 2003; Tatler, et al., 2005; Baddeley & Tatler 2006; Tatler, et al., 2006).

The second right hand term in equation 1 is the Bayesian prior,  $P(\text{saccade})$ . This term describes the probability of saccading to a location irrespective of the visual information at that location, or indeed anywhere in the scene. As such this term will encapsulate any ‘systematic tendencies’ or ‘biases’ in the manner in which we explore scenes with our eyes. Systematic tendencies in oculomotor behaviour can be thought of as regularities that are common across all instances of, and manipulations to, behavioural tasks. Whether or not such systematic biases in how we move our eyes can provide useful insight into predicting fixation selection has not been explored in previous studies of eye guidance. In the present paper we will explore whether understanding these biases can improve our understanding of the moment-to-moment decision about where to target with each saccade.

### **Systematic tendencies in eye guidance**

In contrast to our underdeveloped ability to account for oculomotor selection, in other aspects of motor behaviour, there have been significant recent advances in our ability to model action selection (Körding & Wolpert, 2004). This progress can in part be attributed to recognising that action selection is heavily influenced by the fact that motor behaviours are not all equally likely to be selected. For example, by recording hand movements during daily natural behaviour, it was found that certain combinations of finger movements are far more frequently selected than others (Ingram, Körding, Howard & Wolpert, 2008). In fact 60% of the variance of finger movements during natural behaviour could be described by only the first two principal components. The scale of this result clearly demonstrates that knowledge of this behavioural bias to select certain actions over others is highly informative in our ability to understand and model action selection.

Equivalent approaches have not been used to model visual selection, and our understanding of oculomotor behaviour remains underdeveloped. In the present paper we ask whether a similar approach to that being used to understand other aspects of motor control can be employed to improve current understanding of eye guidance when viewing natural scenes. Two clear questions emerge if we are to ask this: (1) Do oculomotor behavioural biases exist? (2) What is the relative informativeness of any such bias?

The first of these questions is addressed easily: Any survey of the eye movement literature reveals a wealth of support for the notion that there are systematic tendencies to select certain eye movements over others. For example, our oculomotor range is considerable, yet we are not equally likely to make saccades of all possible magnitudes. Instead, across a range of experimental paradigms and environments, saccade magnitudes show a positively-skewed distribution, with a tendency to make small amplitude saccades (e.g., Bahill, Adler & Stark, 1975; Gajewski, Pearson, Mack, Bartlett & Henderson, 2005; Pelz & Canosa, 2001; Tatler et al., 2006). Similarly, we are far from uniform in selecting which direction to execute saccades in (e.g., Bair & O’Keefe, 1998; Lappe, Pekel & Hoffmann, 1998; Lee, Badler & Badler,

2002; Moeller, Kayser, Knecht & König, 2004), with a higher frequency of horizontal saccades than vertical or oblique saccades (but for an exception to this pattern, Bahill et al., 1975, suggested that the majority of saccades made while walking around a real environment were oblique). In recent work we explored the possibility that systematic tendencies may not be limited to the current saccade, but may exhibit sequential dependencies between successive saccades and fixations (Tatler & Vincent, 2008). We found that in many cases the properties of one saccade are influenced by the properties of the fixation and saccade that immediately preceded it (see also Hooge, Over, van Wezel and Frens, 2005; Motter & Belky, 1998; Unema, Panasch, Joos & Velichkovsky, 2005). These studies clearly illustrate that there exist systematic tendencies in the manner in which we move our eyes around a natural scene, and thus such tendencies may offer a previously-untapped source of information about saccade target selection.

The second question will be the focus of the present report. Given that there is evidence for the existence of systematic tendencies in oculomotor control, we ask how informative these are as a component of models of eye guidance. Observing the existence of biases does not mean that they necessarily feature in the moment-to-moment selection of where to fixate. Instead, the observed overall biases may be a consequence of other decision factors, such as where visual information is in the world. As such, any identified biases need not be significant predictors of where each saccade is targeted. If this is the case, we might find that on a saccade-by-saccade basis, visual information predicts fixation selection, but oculomotor biases do not. However, if we find that oculomotor biases are themselves predictive of fixation selection beyond what can be predicted from visual information alone, then the tendency to move our eyes in particular ways can offer an informative component of our understanding of eye guidance.

If oculomotor biases are informative, it will be important to address the question of what these biases reflect: it is not the case that these necessarily reflect purely motoric factors; they may arise from a number of sources. For example, there may be high-level factors that influence these biases – when driving many saccades will be launched horizontally, whereas when batting in cricket vertical saccades will dominate as the ball is bowled (Land, 2006). Whatever the source of these biases, the finding that they are an informative component of fixation selection will be an important contribution to current approaches to modelling eye guidance. We will return to the issue of the possible sources of systematic oculomotor tendencies in the General Discussion.

An established technique for assessing how informative or predictive visual features are in eye guidance is to compare the visual feature content at fixated and control locations (e.g., Parkhurst & Niebur, 2003; Parkhurst, Law and Niebur, 2002; Reinagel & Zador, 1999; Tatler et al., 2005). The logic is that if we can discriminate the feature content at fixated and control locations, then this feature is predictive of where will be fixated. Using techniques such as signal detection, we can further estimate the magnitude of the discrimination between fixated and control locations and use this as an indicator of the extent to which a feature can be informative about fixation selection. In the present report, we use the same principle of determining whether fixated and control locations can be discriminated. However, we extend this technique

to allow us to consider not only the predictive power of visual features, but also that of oculomotor biases. To do this we extract not only the image features at each fixated and control location, but also the amplitude and direction of the saccade immediately preceding each of these real and control fixations (note that this is entirely possible for the control locations because we construct our control locations by sampling from fixations made on other images by the same participant). We evaluated how well we could predict saccade target selection by using a combination of signal detection theory and log-likelihood classifiers to discriminate fixated locations from control locations.

In order to assess the relative informativeness of visual features and oculomotor biases in eye guidance we first assessed how well visual features or oculomotor biases alone could be used to predict fixation selection. We then considered how these factors may combine to predict fixation, either as independent or interactive factors. When testing the ability to predict fixation selection on the basis of visual features, we decided to characterise these features in two different manners. First we used the popular current model of visual salience (Itti & Koch, 2000), which uses a competitive combination of visual feature channels to compute an overall conspicuity map. Second, based on previous work by one of the authors, we characterised visual features simply in terms of the edge intensity derived from the output of oriented Gabor filters, because when interactions between image features are accounted for edge intensity was found to be the maximally informative feature for predicting eye guidance (Baddeley and Tatler, 2006). Thus the present study offers an opportunity not only to assess the relative predictive power of visual features and oculomotor biases, but also to compare the complex salience map model to a simple description of oriented edge information in scenes.

## METHOD

### **Participants**

All 22 participants were naïve to the purpose of the experiment and had normal or corrected-to-normal vision. Each received monetary reward (£5) or course credit toward their undergraduate psychology degree for taking part.

### **Stimuli and procedure**

Each participant viewed 120 colour photographs of natural scenes: 40 indoor, and 80 outdoor scenes (Figure 3). Images were taken using a Nikon D2 digital SLR at a resolution 4 megapixels and later resized to be 1600 x 1200 pixels and represented with 24-bit colour depth.

Images were displayed on a Viewsonic P225f 22" pure flat CRT monitor running at a refresh rate of 100 Hz. At the viewing distance of 57 cm, the images subtended approximately 40° horizontally and 30° vertically in the observer's field of view.

Participants were given no specific task instructions, merely being asked to freely view the images<sup>1</sup>. Each trial was preceded by a fixation marker positioned randomly within 10° of the centre of the screen. Images were presented for five seconds and were followed by a white noise mask.

### **Eye movement recording**

Eye movements were recorded using an SR Research Ltd. EyeLink II eye tracker, sampling pupil position at 500 Hz. Two 9-point grids were used to calibrate and then validate gaze position tracking. If the validation procedure returned a mean spatial accuracy of worse than  $\pm 0.5^\circ$ , the eye tracker was re-calibrated. Eye position data were collected for the eye that produced the better spatial accuracy as determined using the calibration. Saccades and fixations were defined using the saccade detection algorithm supplied by SR Research: Saccades were identified by deflections in eye position in excess of  $0.1^\circ$ , with a minimum velocity of  $30^\circ\text{s}^{-1}$  and a minimum acceleration of  $8000^\circ\text{s}^{-2}$ , maintained for at least 4 ms. We employed a minimum fixation duration of 50 ms. The first fixation in each trial was defined as the first fixation that began after the onset of the scene image. Thus the fixation on the pre-trial fixation marker was not included in the analyses.

### **Defining visual features and saccade characteristics for fixated locations**

*Visual features 1: Edges.*

---

<sup>1</sup> Free viewing is often taken as a task that is free from higher-level 'baggage'. However, this is not the case and free viewing comes with its own set of issues (Tatler et al., 2005). We chose this 'task' for comparability with other recent studies that have evaluated low-level factors in eye guidance.



Given our previous finding that edge information was more predictive of fixation behaviour than luminance or contrast information (Baddeley & Tatler, 2006), in the present paper we chose only to extract edge information at fixated and control locations. Edge information was quantified by convolving images with four oriented odd-phase Gabor patches (oriented at 0°, 45°, 90° and 135°). The absolute values of each orientation map were used in order to capture unsigned difference from the mean. These orientation maps were combined and normalised by subtracting the mean of the combined map and dividing by its standard deviation, as in our previous work (Tatler et al., 2005). Edge maps were constructed over a range of spatial scales, using Gabors with envelope standard deviations from 0.625 to 20 cpd.

In order to extract edge information at fixation, 2°x2° patches were defined, centred around the point of the fixation as derived from the eye tracker record. The maximum edge feature value within this patch was then calculated and used as the value of edge information at fixation (see Figure 2).

### *Visual features 2: Saliency*

We calculated the overall saliency map for each image using the latest version of Itti's saliency algorithm, available at <http://www.saliencytoolbox.net>. We used the default parameter settings for computing our saliency maps. For details of this algorithm see (Walther & Koch, 2006). Just as for edge information, we extracted saliency at fixation by taking the maximum value in a 2°x2° centred with respect to gaze.

### *Saccade characteristics*

Saccade characteristics were also measured for each fixation, extracting the amplitude and direction of the saccade that brought the eye to bear on each fixated location (see Figure 2).

## **Defining visual features and saccade characteristics of control locations**

While we have direct access to fixated image regions from our empirical eye movement recordings, we must also define a set of control locations. There are several ways in which these control locations and their feature and saccade characteristics could be defined. One way to do this would be to uniformly sample values (of edge intensities, for example) within the range of edge intensities present in the image (see Figure 1), however this has the drawback of not accounting for potential correlations in image features over the image. Such correlations may arise from the combination of: any compositional bias in capturing the display images (e.g., a tendency to take photographs with objects of interest in the centre); and any tendency to fixate some parts of images more than others (e.g., the central fixation bias, see Tatler, 2007).

FIGURE 1 ABOUT HERE

Alternatively, these distributions could be created by uniformly and randomly sampling locations in the image (See Figure 1). This approach also has a drawback of

not being a suitable control for the well-known non-random distribution of saccades over an image; there is a strong tendency to fixate the approximate centre of an image (Tatler, 2007).

We therefore employed a method used in our previous studies, which avoids the limitations of the two methods described above (Tatler et al., 2005). The method for creating control distributions of edge intensities, visual salience, saccade magnitudes and saccade directions is shown schematically for one participant viewing one image in Figure 2.

First we compiled a list of x-y locations for fixations, together with the magnitudes and directions of the saccades that immediately preceded each of these fixations. These details were extracted for all fixations on all images excluding the one currently being analysed for this participant. We sampled from this list randomly a number of times equal to the number of actual fixations made on the current image. In doing this, we therefore sampled from actual saccades and fixation locations made by that participant but not on the current image. The control saccade magnitude and direction distributions were built up cumulatively with this procedure over all images and participants. The control feature distributions were constructed by using the sampled x-y locations to extract edge intensity from the current image. This procedure is a suitable way of constructing control image feature distributions because it accounts for the distribution of edge intensities on individual images, while also accounting for any spatial biases of the observer.

FIGURE 2 ABOUT HERE

### **Assessing feature and saccade differences with a log likelihood ratio classifier**

In order to assess how much we can predict about eye movements we constructed a simple classifier to distinguish fixated image patches (F) from control image patches (C). We calculated the performance at distinguishing F from C using a log likelihood threshold criterion with: feature information (either edges or salience); saccade bias information; both feature and saccade bias information (assuming independence); feature and saccade bias (accounting for dependencies between these factors). The corresponding log likelihood ratios are respectively:

$$\log\left(\frac{P(\text{feature} | F)}{P(\text{feature} | C)}\right) \quad (\text{eq. 2})$$

$$\log\left(\frac{P(\text{magnitude, direction} | F)}{P(\text{magnitude, direction} | C)}\right) \quad (\text{eq. 3})$$

$$\log\left(\frac{P(\text{feature, (magnitude, direction)} | F)}{P(\text{feature, (magnitude, direction)} | C)}\right) \quad (\text{eq. 4})$$

$$\log\left(\frac{P(\text{feature,magnitude,direction} | F)}{P(\text{feature,magnitude,direction} | C)}\right) \quad (\text{eq. 5})$$

In the above equations, feature corresponds to edge intensity (see above) of a fixated location, magnitude is the amplitude of the saccade that brought the eye to the fixated location in degrees of visual angle, and direction is the direction in space of the saccade where  $0^\circ$  is rightward,  $-90^\circ$  is upward.

In order to compute the log likelihood ratios for particular patches, we need to define the likelihoods, which are simply probability distributions. We represented these probability distributions using histograms because they are non-parametric (thus containing no assumptions about the shape of the distributions). This approach does have a free parameter (number of histogram bins) per dimension of the probability distribution under consideration.

When producing these histograms, it was important to optimise the number of bins such that the best description of the underlying distributions was provided. We used 10-fold cross validation to evaluate the generalisation performance on the test set of data for histograms constructed using the training set. Increasing the number of bins will always fit the training data better, but as this effectively fits the noise in the data set as well as the signal, the performance on the test set will decrease beyond a certain number of bins. For the final results, we chose 20 bins for each dimension because this corresponded to a clear performance peak when accounting for dependencies between feature and saccade information, and had no impact on other conditions.

Of course building probability distributions from all of the fixated image patches and then testing the performance of this classifier on that very same data could lead to over-fitting, with high performance for these data, but potentially poor generalisation to unseen data. We avoided this issue by using 10-fold cross validation. This splits the dataset up into 10 parts, so we build 10 sets of probability distributions and then report the performance measures on the corresponding 10 test portions of the data. Furthermore, the actual quantity reported is the mean and its 95% confidence intervals across the 10 cross validation test sets as estimated by bootstrap (Efron & Tibshirani, 1993).

For any candidate image location, the log likelihood is calculated by using the appropriate equation from above (2, 3, 4 or 5). If the log likelihood ratio is greater than zero then that patch is more likely to have been fixated (F) than control (C). In this way, over many patches we can calculate percent correct. To clarify, for a particular data point in the cross validation test set, the corresponding log likelihood is determined by linear interpolation from the log likelihood ratios calculated from the training set.

## RESULTS AND DISCUSSION

### Eye movement behaviour

Table 1 shows standard eye movement measures for our dataset. Figure 3 shows example plots of three images from this study overlaid with locations fixated by all participants (left column). We also show the human fixation locations superimposed onto salience maps (middle column) and edge maps (right column) for these scenes.

INSERT TABLE 1 AND FIGURE 3 ABOUT HERE

The Bayesian formulation of how we select the target of our next fixation can effectively be broken down into two constituent parts: that which involves the visual information in the scene:  $P(\text{data}|\text{saccade})/P(\text{data})$ ; and that which involves the systematic oculomotor tendencies:  $P(\text{saccade})$ . From this two-part expression of the problem we can postulate at least four simple hypotheses about how these factors may contribute to eye guidance: (1) selection is based on visual features/salience alone; (2) selection is based on oculomotor tendencies alone, (3) selection involves independent contributions from visual features and oculomotor biases, (4) visual features and oculomotor biases interact to select saccade targets. It is important to note that in all analyses that follow we were testing the ability of a classifier built on a portion of the dataset to generalise to *new, unseen* data. See the Method for further details.

#### 1. Visual features

*Salience.* Using signal detection theory, we found that fixated and control locations could be discriminated. The area under the receiver operator curve (AUC) was 0.565 (with a 95% confidence interval of 0.562-0.568). This result is significantly different from chance discrimination (0.5), but the magnitude of the difference is quite small. A log-likelihood classifier based on visual salience information alone was able to discriminate fixated from control locations with a performance (proportion correct) of 0.554 (95% confidence interval: 0.552-0.557).

*Edges.* For edge information, the AUC was 0.593 (95% confidence interval: 0.590-0.597). For the log-likelihood classifier, edge information provided a discrimination performance (proportion correct) of 0.562 (95% confidence interval: 0.559-0.564).

Our results for edge information alone and for Itti's full salience model result in very similar abilities to discriminate fixated and control locations. These findings are in line with previous studies, which have shown significant effects, but with low magnitude differences (e.g., Einhäuser, Spain & Perona, 2008b; Nyström & Holmqvist, 2008; Tatler et al., 2005)<sup>2</sup>. These results demonstrate that low-level visual information offers some predictive power, but is of limited informativeness in

---

<sup>2</sup> It should be noted that some studies using different methods have reported higher ROC AUC values for Itti's salience map (e.g., Gao, Mahadevan & Vasconcelos, 2008). However, our values fall in the range of previous studies and where in this range we fall does not undermine any comparisons of the *relative* predictive power of salience and motor biases in our dataset.

understanding eye guidance. It should also be noted that the simple edge model performed similarly to the full salience model (in fact it was significantly better for discriminating between fixated and control locations,  $p < .001$ ). This result raises questions about the need for the more complex salience framework. A more parsimonious, yet equally effective description of eye guidance might therefore be that the portion of eye movement behaviour when looking at complex scenes that can be predicted from low-level information can be accounted for in terms of the edge information available in scenes. Previous work from our group has also suggested that when correlations between contrast, luminance and edge information are accounted for, it is edges that provide the best ability to predict fixation selection, and that the apparent predictive power of other features can be attributed to their correlation with the occurrence of edges (Baddeley & Tatler, 2006).

## **2. Oculomotor tendencies**

Next, we assessed how well oculomotor biases alone can be used to predict where human observers fixate. Figure 4 plots the interaction between saccade amplitude and direction in our dataset. Saccades in horizontal directions were more frequent than in vertical directions (see also Tatler & Vincent, 2008). Moreover, horizontal saccades tended to be of larger amplitude than vertical saccades. The plot in Figure 4 can also be seen as effectively a prior probability of saccade targeting in retinocentric space, with the centre of the plot representing the current location of the eye.

FIGURE 4 ABOUT HERE

It should be noted that by assessing the predictive power of oculomotor biases alone, we are effectively testing an extreme and implausible hypothesis: we are effectively evaluating a model of fixation selection that is *blind* to visual information in the scene. If low-level visual salience is a prominent factor in selecting where to fixate we would expect that this extreme behavioural bias hypothesis should not perform as well as the salience or edge models. However, a log-likelihood classifier was able to discriminate fixated and control locations on the basis of oculomotor biases alone with a performance of 0.648 (95% confidence interval: 0.645-0.650). Therefore, feature-based accounts of eye guidance are out-performed when we know only about oculomotor biases and know *nothing* about image features. This striking result indicates that biases in how we move our eyes can be highly informative about the locations in the world that we select to look at. This result alone poses serious challenges to the existing feature- and salience-based frameworks for explaining eye movement behaviour. That we can do at least as well, if not considerably better, at predicting fixation selection using only knowledge of how we tend to move the eyes, underlines the limited explanatory power of existing low-level frameworks.

## **3. Independent combination of visual features and oculomotor biases**

Of course, both of the hypotheses tested above are straw men: No-one really suggests that eye guidance operates solely on the basis of low-level image properties. Similarly, oculomotor biases alone (with no contribution from visual input!) could not be the sole factor. Hence in this and the following section we ask whether we can use

oculomotor biases alongside visual feature information in order to increase our ability to account for where people fixate.

Performances (in terms of proportion correct) for classifiers based upon independent contributions of oculomotor biases and either salience (0.687, 95% CI: 0.685-0.690) or edges (0.691, 95% CI: 0.685-0.691) were better than for either oculomotor biases or visual features alone. Thus we can see that accounting for both visual information and oculomotor characteristics improves our ability to discriminate locations that were fixated from control location. We can use this result to suggest that using a combination of knowledge of what visual information is present in the scene and how humans tend to move their eyes is a useful and informative way of framing eye guidance.

#### **4. Interactive combination of visual features and oculomotor biases**

Finally, we considered the performance of a classifier based upon the interaction between oculomotor tendencies and visual features. For both the salience- and edge-based versions, we see very high levels of classification performance: For the interaction between salience and oculomotor characteristics, discrimination performance was 0.800 (95% confidence interval: 0.710-0.894). For the interaction between edge and oculomotor characteristics, discrimination performance was 0.822 (95% confidence interval: 0.785-0.862). Thus there is clear evidence of a strong interaction: i.e. if any feature selection is occurring, then its nature varies over the visual field.

For ease of comparison we depict the performances of our various log-likelihood classifiers graphically in Figure 5. AUC values are also included for the salience- and edge-only classifiers in order to allow comparison to previous studies.

FIGURE 5 ABOUT HERE

## GENERAL DISCUSSION

Like a large number of researchers before us, we set out to ask how well we can explain where people fixate in images of natural scenes. However, unlike many previous studies, we chose to ask this question from a Bayesian point of view. Simply by posing the question of how we select the target of each saccade in this way, we immediately identified a previously neglected component of the moment-to-moment decision process: the prior probability of saccading to a location irrespective of the visual information present. This prior probability is interesting in two respects: First it can be seen as capturing any systematic biases that exist in how we move our eyes around the visual environment. Second, we found it to be a highly informative component of eye guidance.

The ability to predict fixation selection from salience or visual features alone in our dataset was in line with a variety of recent studies that have shown significant but small predictive ability of low-level factors (e.g., Einhäuser et al., 2008b; Nyström, & Holmqvist, 2008). Our finding that oculomotor biases alone were able to predict fixation selection better than salience or edge information clearly demonstrates that this is a useful source of information about where people will select as the target of each saccade. A visual priority map such as Itti's salience model is thus a much-oversimplified description of fixation selection. We have shown that by adding to this knowledge of where the eye currently is and how it tends to move, we can arrive at a surprisingly good description of what will be selected as the target of each new fixation.

Our Bayesian framing of the moment-to-moment relocation of gaze has thus provided a very simple account whereby fixation selection involves the combination of both the (low-level) visual information in the scene and the tendency to move the eyes in particular ways. Of course, expanding the account to include specific higher-level cues will be crucial in further understanding fixation selection. Importantly, extrapolation to any number or any form of cues is achievable within the same simple Bayesian explanatory approach.

### **The origins and implications of oculomotor biases in eye guidance**

Given the predictive power of oculomotor biases in our dataset it is important to consider these biases in more detail. As we stated in the Introduction, the observation of oculomotor biases does not reveal their origin. Certainly, we do not suggest that these biases are purely motoric in origin. They may arise from a number of factors, from 'low-level' biomechanics, to our learnt knowledge of the structure of the world and the distribution of objects of interest. We shall now consider some possible sources that may contribute to the observed biases.

#### *1. Biomechanical factors*

Oblique movements are executed by coordinating horizontal and vertical muscular activity (Becker & Jurgens, 1990) and may have lower efficiency (Pitzalis & Di Russo, 2001) than saccades that only involve horizontal or vertical components. It is plausible that there are costs associated with coordinating horizontal and vertical components of an eye movement such that movements accomplished primarily by

only one set of oculomotor muscles are favoured (Smit, Van Gisbergen & Cools, 1987; Viviani, Berthoz & Tracey, 1977). A tendency to make smaller amplitude saccades in favour of larger amplitude saccades could again be argued on the basis of metabolic costs: the energy required to create greater tension in the muscles over a longer period of time for large saccades than for small saccades would mean that larger saccades were more costly. However, metabolic costs are unlikely to be a major factor in eye movement control because the inertia of the eyeball is very low (Carpenter, 1988; Robinson, 1964).

### *2. Saccade flight time and landing accuracy*

Saccade targeting precision is likely to be influenced by peripheral acuity limits and ‘crowding’ effects (e.g., Bouma, 1970; Stuart & Burian, 1962) and indeed smaller amplitude saccades tend to be more accurate (e.g., Becker, 1991). Thus targeting accuracy could favour small saccades. Small amplitude saccades also minimise flight time (e.g., Carpenter, 1988; Robinson, 1964; Collewijn, Erkelens & Steinman, 1988). As such, small amplitude saccades will minimise the time that vision is disrupted (e.g., Burr, Morrone & Ross, 1994; Matin, 1974; Volkman, 1976).

### *3. Uncertainty*

An emerging theme in eye guidance is that saccades may reduce local uncertainty (Najemnik & Geisler, 2005; Renninger, Vergheese & Coughlan, 2007; Sprague, Ballard & Robinson, 2007). The decline in sampling density with eccentricity in the retina is radially asymmetric, declining more rapidly with vertical eccentricity than with horizontal eccentricity (Curcio, Sloan, Kalina & Hendrickson, 1990). This means that uncertainty increases more rapidly with eccentricity in vertical directions. At first this may appear to favour the opposite of our findings: that more vertical saccades should be observed. However, Najemnik & Geisler (2008) showed that an ideal observer that incorporated this asymmetry in visual sampling produced distributions of saccade direction and amplitude in a search task that closely match our observed distributions (their Figure 5). Thus it would appear that the principle of using saccade to reduce uncertainty is consistent with our findings.

### *4. Distribution of objects of interest in the environment*

Properties of the visual environment itself may result in a spatially variant utility function for saccade behaviour. Studies of natural scene statistics show that there is often most power in the horizontal directions, followed by vertical and finally oblique directions (Torralba & Oliva, 2003). It is interesting to compare this to the relative frequency of saccade directions we observed (Figure 4): perhaps the abundance of horizontal and vertical form in natural scenes influences saccade targeting.

Not only are image statistics non-uniform, but they are also non-uniformly distributed in space. Baddeley (1997) showed that, from the autocorrelation function of natural scenes, any two features are *on average* more similar the closer they are in space. Correspondingly, regardless of where we fixate, features further from the fovea will become on average more different to the location currently fixated. Not only do correlations decrease with eccentricity, but the decrease is more rapid in the vertical direction than in the horizontal direction from fixation (Baddeley, 1997). While the correlations in image statistics follow the same pattern over retinotopic space as does the distribution of oculomotor biases that we have described, it is hard to argue that



these correlations might underlie our proposed saccade utility function. Indeed most saccade targeting models assume the opposite: salience based models (e.g., Itti & Koch, 2000; Parkhurst et al., 2002) identify locations with maximal difference from their surroundings as likely candidates for saccade targeting.

It may be that the biases are shaped at a much higher level. In natural tasks, objects of interest to us in completing the task will not be uniformly distributed around us. For example, when making tea in a kitchen (e.g., Land et al., 1999) the objects of interest will tend to lie in a horizontal envelope around us; on or near work surfaces, rather than on the floor or ceiling. Similar distributions of objects will occur in many real tasks, although of course there are exceptions. Could it be that the shape of the saccadic biases that we have measured reflects this asymmetric distribution of objects in the world around fixation? The idea that the distribution of photoreceptors may be optimised to object distributions has been suggested before (Lewis, Garcia & Li, 2003). Lewis and colleagues used sampling theory to suggest that the distribution of objects in the visual environment was very similar to the distribution of cones in the human retina.

#### *5. Task parameters*

Finally, it is possible that different behavioural tasks favour different strategies for looking. That is, biases in oculomotor behaviour may arise from overall behavioural strategies. Anecdotally, this seems plausible given it is easy to generate thought experiments that would favour particular sets of eye movements over others. There is also empirical evidence that saccade metrics vary between tasks (Rayner, Li, Williams, Cave & Well, 2007; Steinman, 2003; Tatler et al., 2006) and strategies (Gilchrist & Harvey, 2006). However, there has yet to be a systematic exploration of this possibility.

#### **Ecological validity of the described oculomotor biases**

In the present study, data were collected with participants seated in a chair, viewing images presented on a computer screen and using a head-mounted eye tracker. As such, the experimental setting differs in a number of ways from natural behaviour and therefore caution must be exercised when extrapolating conclusions drawn about the nature of the prior knowledge and about its role in saccade behaviour to real world settings. Two factors are particularly important when attempting to evaluate the ecological validity of our findings.

First, photographic scenes are in many ways removed from natural environments, in terms of the variation in luminance, depth cues and lack of dynamics. Furthermore, they are often compositionally biased by implicit or explicit tendencies of the photographer to place objects of interest near the centre of the scenes. Placing the scene within the bounds of a computer monitor is also likely to influence inspection behaviour. The 40° by 30° monitor is much smaller than the approximately 180° horizontally by about 80° vertically human binocular field of view. Under natural conditions gaze shifts (incorporating eye, head and body rotation) can be up to 180° (Land et al, 1999; Land, 2004). Certainly, saccade magnitude distributions are influenced by the extent of the image viewed: while the long-tailed distribution remains, it scales to the size of the image observed (von Wartburg et al, 2007). Under

natural viewing conditions, saccade amplitudes are much greater than when viewing a computer screen, with mean saccade amplitudes of up to 20° (Land et al, 1999). The screen not only influences the amplitudes of saccades but also introduces a strong bias to look near the centre of the screen, irrespective of the distribution of visual features in the scene (Tatler, 2007). Thus, the nature of the biases described in our study (Figure 4) may reflect aspects of the artificial nature of viewing images on a computer screen. However, we feel that this does not weaken our findings because our aim, like that of many other contemporary groups, was to consider the factors that allow us to explain eye movement behaviour in our experimental setting. Furthermore, even if the saccade biases look different during natural behaviour, they may still provide a highly informative component of the decision about where to target the next saccade.

Second, data were only collected for a single behavioural task in the present study. Theoretically, if we are to describe oculomotor biases fully we should marginalise fixation data over *all* possible behaviours. By limiting ourselves to only considering free viewing of scenes, the form of the saccadic biases may be of limited generalisability to other behavioural tasks.

### **Improving our description of oculomotor biases**

It should be noted that our description of systematic tendencies in saccade generation is very simple: we describe this only in terms of the amplitude and direction of each saccade. Describing oculomotor biases in this way effectively treats each saccade as an independent event, yet we know that this is not the case: previous studies have shown that the amplitude and direction of each saccade can be heavily influenced by the amplitude and direction preceding saccade and the duration of the fixation from which the saccade is launched (e.g., Tatler & Vincent, 2008; Unema et al., 2005). As such, a more comprehensive description of systematic oculomotor biases should incorporate these sequential effects into any model of eye guidance. Not only do we know that there are sequential dependencies between successive eye movements, but we also know that saccade behaviour changes over viewing time (Tatler et al., 2005). Understanding when an eye movement is launched is likely to offer further insights into behavioural biases and the moment-to-moment decision about where to fixate.

### **What is the future for salience models?**

We feel that the research field has largely reached somewhat of an impasse in salience-based approaches to eye guidance. There abound reports of specific failures of the basic model to account for certain behavioural tasks. Certainly consensus in the community is now that any low-level salience-type scheme will be quite limited in its ability to account for the complexities of eye movement behaviour. Our result that a visually blind model based only on biases in how we initiate saccades could dramatically outperform salience and edge models should in itself perhaps collectively motivate us to consider other lines of enquiry. As such, the challenge in this field must be to find the right way to move on from this impasse. We feel that an understanding of incorporating knowledge of *how* we tend to move our eyes will benefit emerging probabilistic models of eye guidance that attempt to incorporate higher-level factors (Ehinger et al., this issue; Kanan et al., this issue).

**ACKNOWLEDGEMENTS**

We thank Jim Brockmole, Aude Oliva and Simon Liversedge for their comments on the submitted version of this manuscript. We also thank Roland Baddeley, Mark Bennett, Tim Dixon, Martin Fischer, Gesche Huebner, Casimir Ludwig, Wayne Murray and Nick Wade for their comments and guidance during the preparation of this work.

REFERENCES

- Baddeley, R. (1997). The correlational structure of natural images and the calibration of spatial representations. *Cognitive Science*, 21(3), 351-371.
- Baddeley, R. J., & Tatler, B. W. (2006). High frequency edges (but not contrast) predict where we fixate: a Bayesian system identification analysis. *Vision Research*, 46, 2824-2833.
- Bahill, A. T., Adler, D., & Stark, L. (1975). Most Naturally Occurring Human Saccades Have Magnitudes of 15 Degrees or Less. *Investigative Ophthalmology*, 14(6), 468-469.
- Bair, W., & O'Keefe, L. P. (1998). The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Visual Neuroscience*, 15(4), 779-786.
- Ballard, D. H., Hayhoe, M. M., Li, F., Whitehead, S. D., Frisby, J. P., Taylor, J. G., et al. (1992). Hand Eye Coordination During Sequential Tasks. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 337(1281), 331-339.
- Becker, W., & Jurgens, R. (1990). Human Oblique Saccades - Quantitative-Analysis of the Relation between Horizontal and Vertical Components. *Vision Research*, 30(6), 893-920.
- Bouma, H. (1970). Interaction Effects in Parafoveal Letter Recognition. *Nature*, 226(5241), 177-178.
- Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective Suppression of the Magnocellular Visual Pathway During Saccadic Eye-Movements. *Nature*, 371(6497), 511-513.
- Buswell, G. T. (1935). *How People Look at Pictures: A Study of the Psychology of Perception in Art*. Chicago: University of Chicago Press.
- Carpenter, R. H. S. (1988). *Movements of the eyes* (2nd ed.). London: Pion.
- Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1988). Binocular Coordination of Human Vertical Saccadic Eye-Movements. *Journal of Physiology-London*, 404, 183-197.
- Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human Photoreceptor Topography. *Journal of Comparative Neurology*, 292(4), 497-523.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and visual similarity. *Psychological Review*, 96, 433-458.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall.
- Einhäuser, W., Rutishauser, U., & Koch, C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision*, 8(2), 2: 1-19.
- Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, 8(14), 18: 11-26.
- Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, 8(2).
- Gajewski, D. A., Pearson, A. M., Mack, M. L., Bartlett, F. N., & Henderson, J. M. (2005). Human gaze control in real world search. In *Attention and Performance in Computational Vision* (Vol. 3368, pp. 83-99).

- Gao, D., Mahadevan, V., & Vasconcelos, N. (2008). On the plausibility of the discriminant center-surround hypothesis for visual saliency. *Journal of Vision*, 8(7), 13: 11-18.
- Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scan paths in visual search. *Visual Cognition*, 14(4-8), 704-715.
- 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., Brockmole, J. R., Castelano, M. S., & Mack, M. L. (2007). Visual saliency does not account for eye movements during search in real-world scenes. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537-562). Oxford: Elsevier.
- Hooge, I.T.C., Over, E.A.B., van Wezel, R.J.A. and Frens, M.A. (2005). Inhibition of return is not a foraging facilitator in saccadic search and free viewing, *Vision Research*, 45, 1901–1908.
- Ingram, K. N., Körding, K. P., Howard, I. S., & Wolpert, D. M. (2008). The statistics of natural hand movements. *Experimental Brain Research*, 188(2), 223-236.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10-12), 1489-1506.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254-1259.
- Koch, C., & Ullman, S. (1985). Shifts in Selective Visual-Attention - Towards the Underlying Neural Circuitry. *Human Neurobiology*, 4(4), 219-227.
- Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244-247.
- Krieger, G., Rentschler, I., Hauske, G., Schill, K., & Zetzsche, C. (2000). Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics. *Spatial Vision*, 13(2-3), 201-214.
- Land, M. F. (2004). The coordination of rotations of the eyes, head and trunk in saccadic turns produced in natural situations. *Experimental Brain Research*, 159(2), 151-160.
- Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, 25(3), 296-324.
- Land, M. F., & Hayhoe, M. M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25-26), 3559-3565.
- Land, M. F., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, 28(11), 1311-1328.
- Lappe, M., Pekel, M., & Hoffmann, K. P. (1998). Optokinetic eye movements elicited by radial optic flow in the macaque monkey. *Journal of Neurophysiology*, 79(3), 1461-1480.
- Lee, S. P., Badler, J. B., & Badler, N. I. (2002). Eyes alive. *Acm Transactions on Graphics*, 21(3), 637-644.
- Lewis, A., Garcia, R., & Zhaoping, L. (2003). The distribution of visual objects on the retina: connecting eye movements and cone distributions. *Journal of Vision*, 3(11), 893-905.
- Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1997). Fixation sequences made during visual examination of briefly presented 2D images. *Spatial Vision*, 11(2), 157-178.

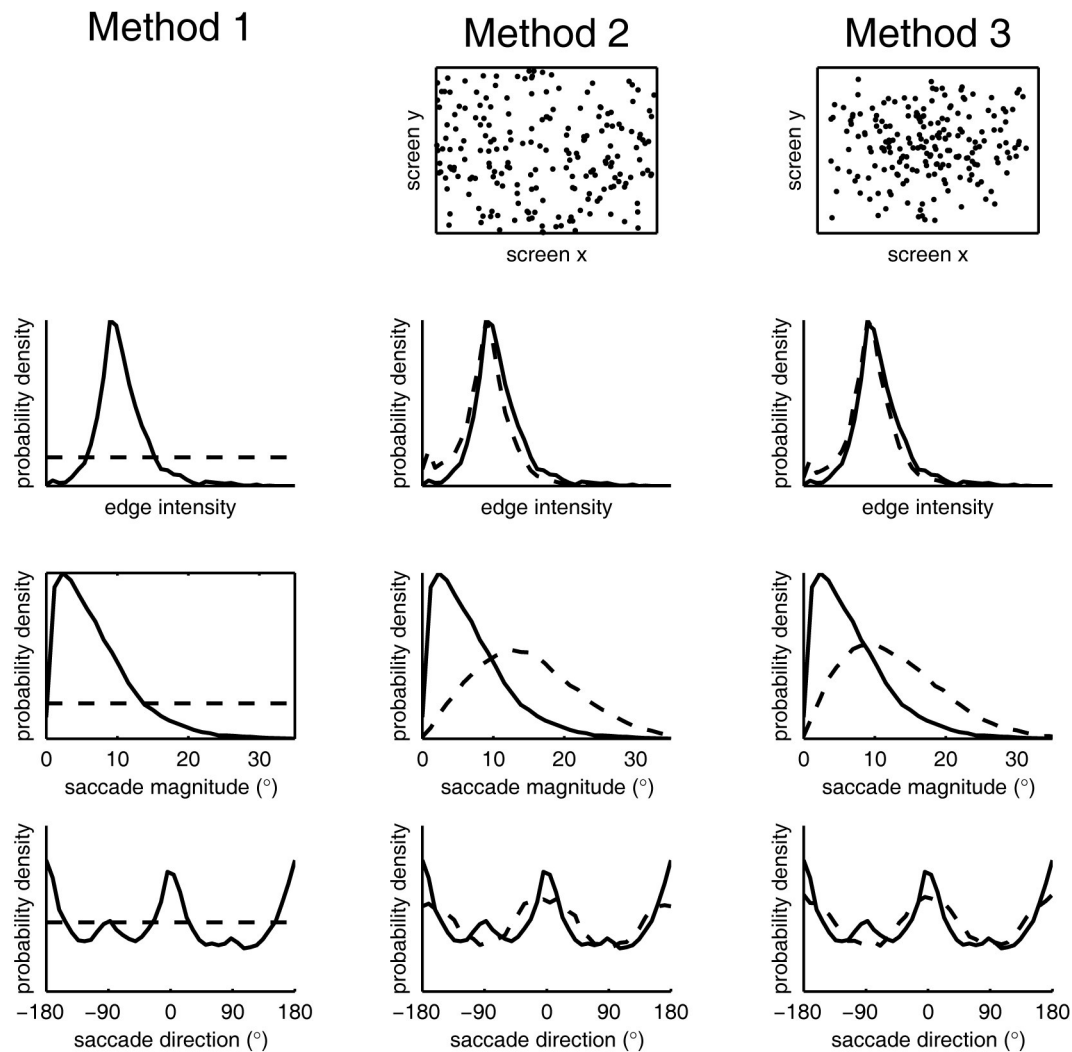
- Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, 81, 899-917.
- Moeller, G. U., Kayser, C., Knecht, F., & Konig, P. (2004). Interactions between eye movement systems in cats and humans. *Experimental Brain Research*, 157(2), 215-224.
- Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research*, 38(12), 1805-1815.
- Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, 8(3), 4: 1-14.
- Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, 45(2), 205-231.
- Nyström, M. & Holmqvist, K. (2008). Semantic override of low-level features in image viewing – both initially and overall. *Journal of Eye Movement Research*, 2(2): 2, 1-11
- Parkhurst, D. J., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, 16(2), 125-154.
- Parkhurst, D. J., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42(1), 107-123.
- Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, 41(25-26), 3587-3596.
- Pitzalis, S., & Di Russo, F. (2001). Spatial anisotropy of saccadic latency in normal subjects and brain-damaged patients. *Cortex*, 37(4), 475-492.
- Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research*, 47(21), 2714-2726.
- Reinagel, P., & Zador, A. M. (1999). Natural scene statistics at the centre of gaze. *Network-Computation in Neural Systems*, 10(4), 341-350.
- Robinson, D. A. (1964). The mechanisms of human smooth pursuit eye movement. *Journal of Physiology*, 180(3), 569-591.
- Smit, A. C., Vangisbergen, J. A. M., & Cools, A. R. (1987). A Parametric Analysis of Human Saccades in Different Experimental Paradigms. *Vision Research*, 27(10), 1745-1762.
- Sprague, N., Ballard, D., & Robinson, A. (2007). Modeling embodied visual behaviors. *ACM Transactions on Applied Perception*, 4.
- Steinman, R. M. (2003). Gaze control under natural conditions. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences*. Cambridge, MA: MIT Press.
- Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty. Its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology*, 53, 471-477.
- Tatler, B. W. (2007). The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14), 4, 1-17.
- Tatler, B. W., & Vincent, B. T. (2008). Systematic tendencies in scene viewing. *Journal of Eye Movement Research*, 2(2): 5, 1-18
- Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, 45(5), 643-659.
- Tatler, B. W., Baddeley, R. J., & Vincent, B. T. (2006). The long and the short of it: spatial statistics at fixation vary with saccade amplitude and task. *Vision Research*, 46, 1857-1862.

- Torralba, A. (2003). Contextual priming for object detection. *International Journal of Computer Vision*, 53(2), 169-191.
- Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network-Computation in Neural Systems*, 14(3), 391-412.
- 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(4), 766-786.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruity influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, 59(11), 1931-1949.
- Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, 18(3), 321-342.
- Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, 12(3), 473-494
- Vincent, B. T., Troscianko, T., & Gilchrist, I. D. (2007). Investigating a space-variant weighted salience account of visual selection. *Vision Research*, 47(13), 1809-1820.
- Viviani, P., Berthoz, A., & Tracey, D. (1977). Curvature of Oblique Saccades. *Vision Research*, 17(5), 661-664.
- Volkman, F. C. (1976). Saccadic Suppression: A Brief Review. In R. A. Monty & J. W. Senders (Eds.), *Eye Movements and Psychological Processes* (pp. 73-83). Hillsdale, NJ: Lawrence Erlbaum Associates.
- von Wartburg, R., Wurtz, P., Pflugshaupt, T., Nyffeler, T., Lüthi, M., & Müri, R. M. (2007). Size matters: Saccades during scene perception. *Perception*, 36(3), 355-365.
- Walther, D. & Koch, C. (2006), Modeling attention to salient proto-objects. *Neural Networks* 19, 1395-1407.
- Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, 9(1), 33-39.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6), 495-501.

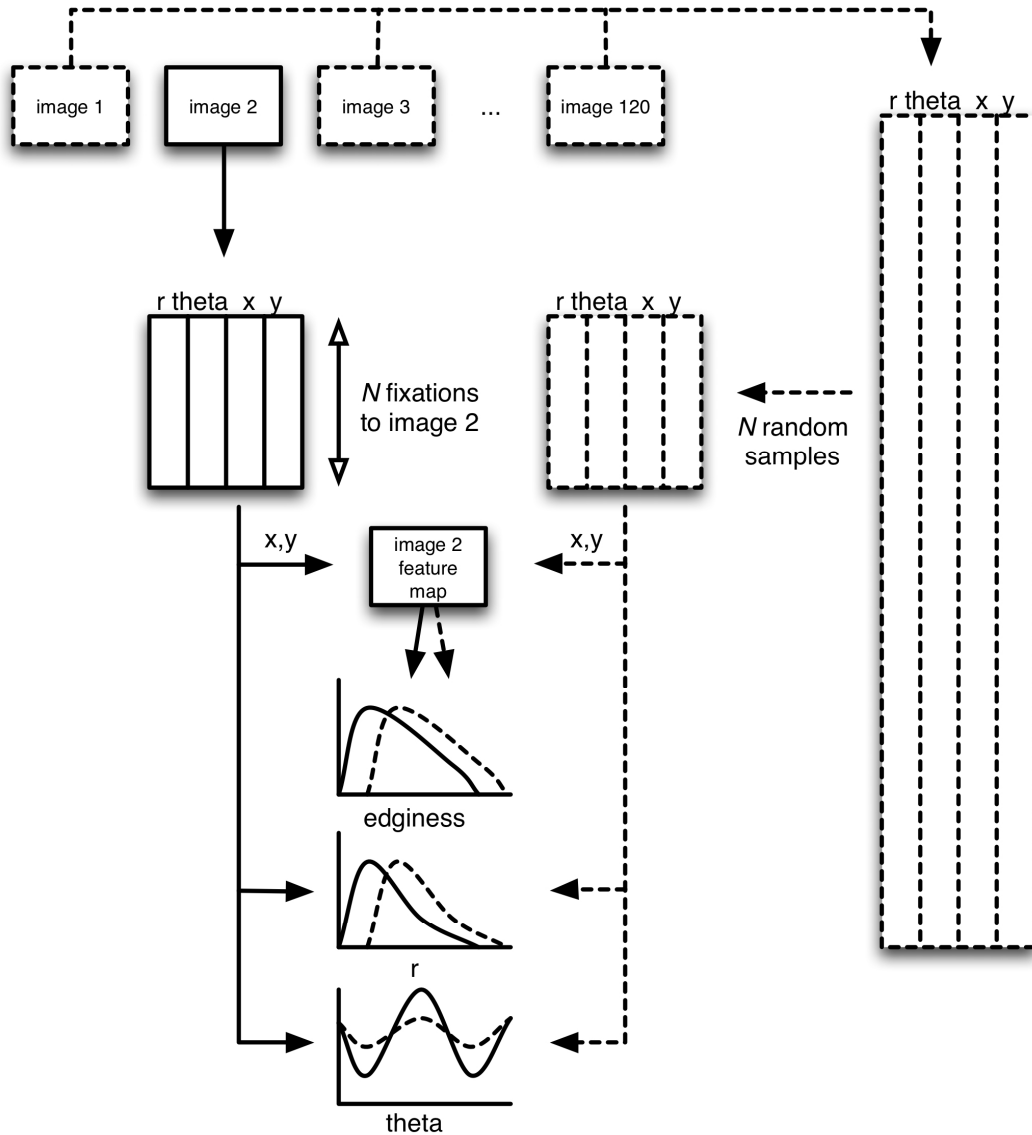
	Mean	Median	Standard deviation
Number of fixations per image (5 s viewing period)	15.0	14.9	1.33
Fixation duration (ms)	258	250	26.3
Saccade amplitude (degrees)	7.03	7.22	0.982

TABLE 1. Summary of eye movement measures between participants. All measures are reported to 3sf.





**FIGURE 1. Methods for sampling non-fixated image locations.** Distributions are shown for edge intensities, saccade amplitudes and saccade directions for fixated (solid lines) and control (dashed lines) locations. In Method 1 a uniform distribution over each factor for the non-fixated locations is defined. In Method 2 we sample uniformly in space and construct the probability distributions of control locations from these locations. In Method 3 control distributions of locations were constructed by randomly sampling from the set of all fixations made by each individual observer on all images, excluding the current image. These coordinates were used to extract image feature information from the current image.



**FIGURE 2. Schematic of how fixated (solid) and control (dashed) distributions are calculated.** This example demonstrates the procedure for one particular participant viewing one particular image (image 2) and is repeated over all images and participants. Fixated edge-intensities are simply drawn from the edge intensity map (of that image) at locations that subjects fixated. Distributions of saccade magnitudes and directions to actually fixated locations are simply calculated from the eye tracker data. Similar distributions for control locations are calculated by randomly sampling from a set of all actual fixations made by a participant to all images excluding image 2. Saccade magnitude and directions come straight from this sample; edge-intensities (and salience) are drawn from these locations but from the edge-intensity (or salience) map from image 2.

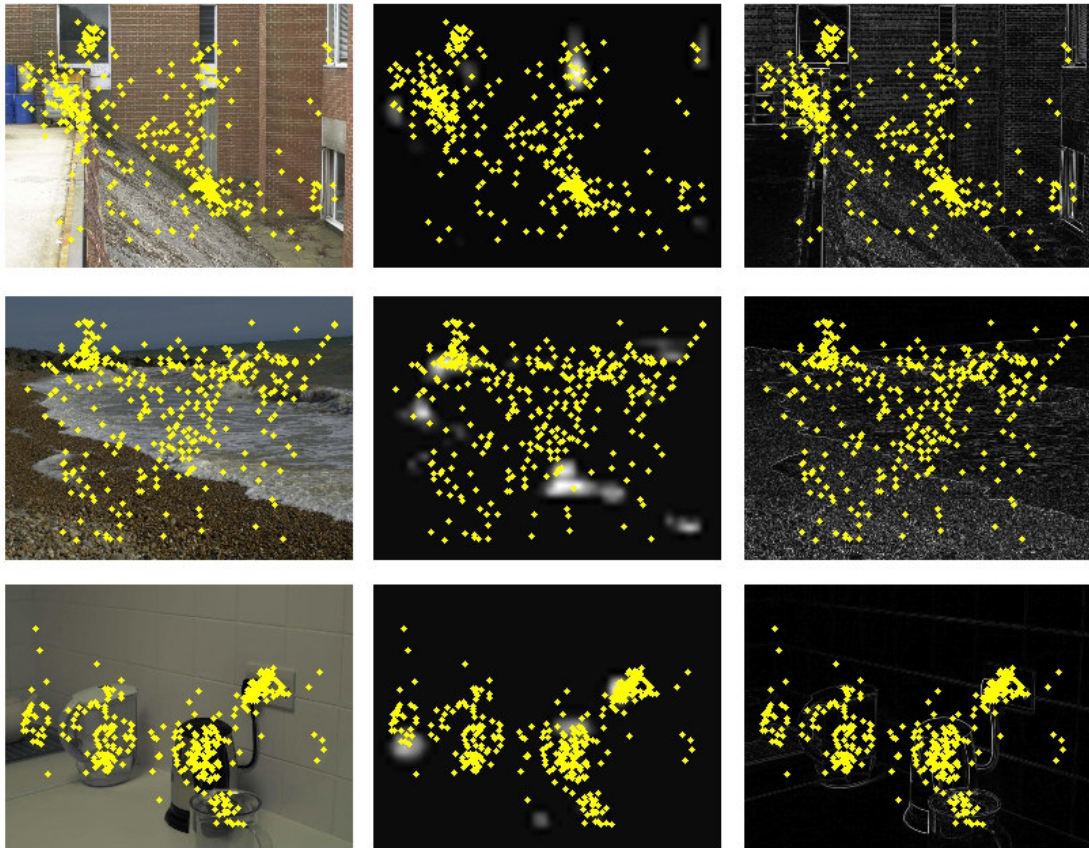
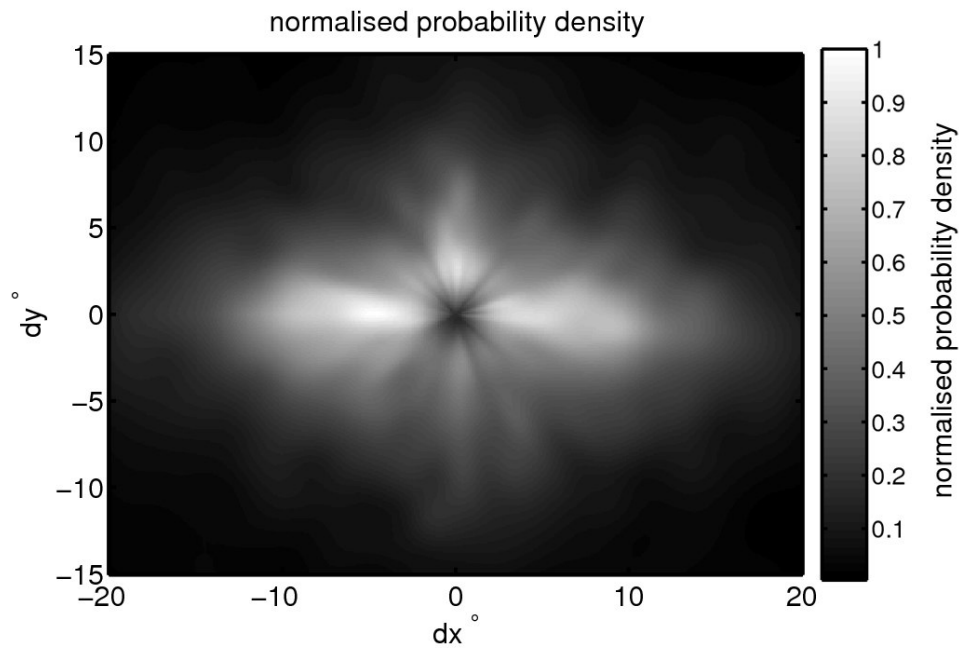
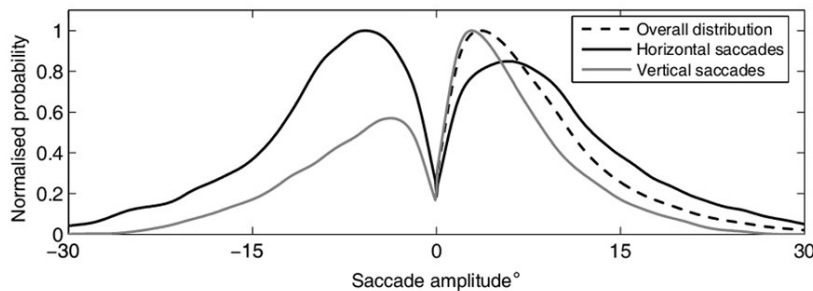


FIGURE 3. Three sample images from our study, with fixation positions overlaid (small dots). For comparison, saliency maps (middle column) and edge maps (right column) are shown for each image, onto which the fixation locations are also overlaid. Fixations are from all participants in the study.

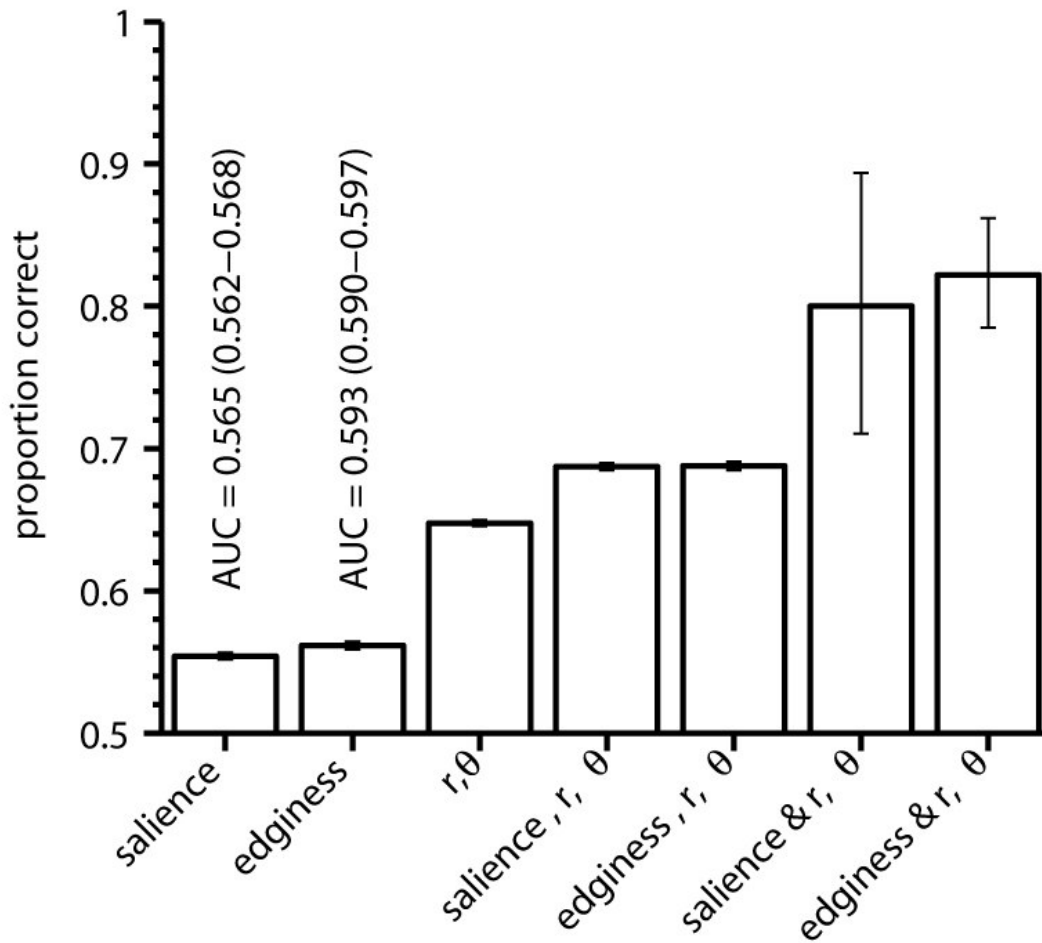
**(A)**



**(B)**



**FIGURE 4. Saccadic biases.** (A) The probability density of making saccades of particular magnitudes and directions. There is a clear bias to make saccades in horizontal directions more frequently than vertical directions, and to make vertical saccades more frequently than saccades in oblique directions. It is interesting to note that saccades were more frequently made in an upward direction than in a downward direction. (B) Cross sections through the distribution plotted in (A). The dashed line shows the distribution of saccade amplitudes for saccades made in all directions. There were more longer saccades and fewer small amplitude saccades for horizontal saccades (solid black line) than for the overall distribution (dashed line). In contrast, vertical saccades (solid grey line) were more frequently of small amplitude than either the overall distribution, or horizontal saccades.



**FIGURE 5. Performance of the classifiers.** Performances, as proportion of correct classifications for each of the 7 classifiers tested in the present study. From left to right, the classifiers tested the ability to correctly classify control and fixated locations on the basis of (1) Itti’s saliency model alone, (2) edge information alone, (3) saccade bias information (saccade amplitude,  $r$ , and saccade direction,  $\theta$ ) alone, (4) the independent combination of saliency and oculomotor biases, (5) the independent combination of edges and oculomotor biases, (6) the interaction between saliency and motor biases, and (7) the interaction between edges and oculomotor biases. Error bars indicate 95% confidence intervals across the 10 cross validation test sets, estimated by bootstrap. For comparison to other studies we plot the area under the receiver operator curve (AUC) for saliency and edge information-based fixation selection.