The saccadic flow baseline: Accounting for image-independent biases in fixation behavior

Alasdair D. F. Clarke

Department of Psychology, University of Essex, Colchester, UK

Matthew J. Stainer

School of Psychology, University of Aberdeen, Aberdeen, UK

Benjamin W. Tatler

School of Psychology, University of Aberdeen, Aberdeen, UK

Amelia R. Hunt

School of Psychology, University of Aberdeen, Aberdeen, UK

Much effort has been made to explain eye guidance during natural scene viewing. However, a substantial component of fixation placement appears to be a set of consistent biases in eye movement behavior. We introduce the concept of *saccadic flow*, a generalization of the central bias that describes the image-independent conditional probability of making a saccade to (x_{i+1}, y_{i+1}), given a fixation at (x_i, y_i). We suggest that saccadic flow can be a useful prior when carrying out analyses of fixation locations, and can be used as a submodule in models of eye movements during scene viewing. We demonstrate the utility of this idea by presenting bias-weighted gaze landscapes, and show that there is a link between the likelihood of a saccade under the flow model, and the salience of the following fixation. We also present a minor improvement to our central bias model (based on using a multivariate truncated Gaussian), and investigate the leftwards and coarse-to-fine biases in scene viewing.

the content of the scene or the task being performed (Foulsham & Kingstone, 2010; Tatler & Vincent, 2009), such as a strong tendency to fixate near to the center of images (Canosa, Pelz, Mennie, & Peak, 2003; Stainer, Scott-Brown, & Tatler, 2013; Tatler, 2007). If we are to gain a complete understanding of the factors that govern how we sample information, we must build models of eye guidance on the framework of these underlying biases, using them as a baseline against which to compare effects of the scene, task, image properties, and individual differences.

Eye movement heuristics

One influential model of eye movements of the last decade is the optimal search model (Najemnik & Geisler, 2008), which posits that human saccadic behavior during visual search is consistent with predictions made by an ideal observer. The number of fixations human observers needed to make to find the target was closely matched by the ideal observer model, in which successive fixations were selected based on reducing uncertainty about the target's location, taking into account search history and target visibility across the visual field. The efficiency of human search (at least, in search for a Gabor patch hidden in $1/f$-noise) suggests this as a plausible mechanism for selecting fixations during search. Further evidence for an optimal strategy comes from Ma, Navalpakkam, Beck, Van Den Berg, and Pouget (2011) who found that human observers are near-optimal in a visual search task with line segments, and presented a neural

Introduction

The human fovea provides a small window of high acuity vision to the world, and the locations that we select to view through this window can tell us how we seek the information necessary to complete the task we are currently undertaking. Fixation locations are selected based on a combination of low-level factors such as visual salience (Borji & Itti, 2013) and high-level factors (Buswell, 1935; Land & Hayhoe, 2001; Yarbus, 1967). However, there are also strong observable biases in eye movements that are independent of
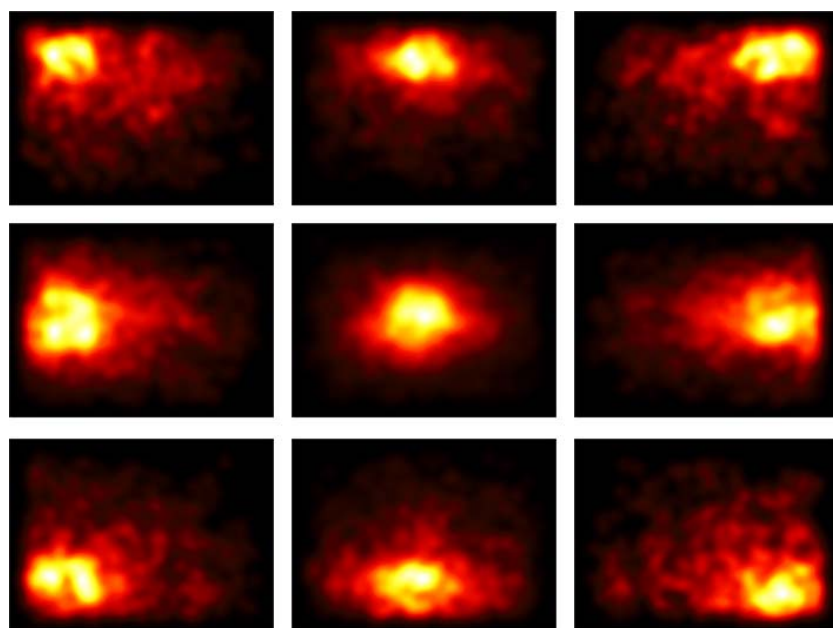
Figure 1. Saccade landing positions from fixations that were in different sections of the screen. Data from each plot has been separated into fixations in nine spatial bins, with the screen being divided into thirds in both horizontal and vertical aspects.

network implementation of near-optimal search based on probabilistic population coding.

While this modeling framework is attractive, there are several issues. The computations driving each fixation are complex, and depend on a fairly precise representation of one's own acuity over the visual field for a wide range of possible target/background combinations. One might therefore question the assumption that these computations are undertaken to determine the location of each of the 3–4 fixations made on average every second during visual search. More importantly, Morvan and Maloney (2012) demonstrated that human observers are not able to use information about visual sensitivity in the periphery to rationally plan even a single saccade to the optimal location in a target discrimination task. In their experiment, the observer simply has to select a location from which to detect a target that can appear with equal probability in one of two possible locations. If the locations are relatively close together, a location in between will maximize the probability of detecting a target appearing in either location. When the targets are too far apart to reliably detect the target from a point equidistant between them, the rational strategy is to look directly at one of the two possible target locations. Inconsistent with optimal viewing strategies, however, the observers did not systematically modify their choice about where to fixate according to the distance between the possible target locations. This striking failure of optimality has recently been replicated in a larger sample and generalized to other decisions in addition to eye movements (Clarke & Hunt, 2016. Further work from Nowakowska, Clarke,

and Hunt (2017) used a simple visual search display (an array of line segments) in which the target was easy to find (pop-out) when present in one half of the display, and hard to find when in the other. The optimal search strategy in this scenario is to search the difficult half of the display: If the target was present in the easy half, the observer would be able to find it using peripheral vision. Whereas a minority of observers followed this search strategy, the majority exhibited large deviations, searching both halves equally, or even fixating the easy half and neglecting the difficult half of the display. These results are consistent with Morvan and Maloney's explanation for the contradiction between their results and the predictions of Najemnik and Geisler's (2008) optimal model: They propose that heuristics guide saccade planning. Heuristics include basic oculomotor biases such as a tendency to make saccades of particular amplitudes, and/or to particular regions of a display, or in particular sequences, depending on the current task.

This idea has recently been formalized in a model by Clarke, Green, Chantler, and Hunt (2016), who demonstrate that a stochastic search model based on a memoryless random walk can find a target in noise in a similar number of fixations to human observers. The key component of this model was the use of the empirical distribution of saccades: For each saccade the model randomly samples a saccade from distributions estimating the likelihood a human observer made a saccade from (x_{i+1}, y_{i+1}) to (x_i, y_i). It is clear from Figure 1 that the distribution of saccade end points varies considerably depending on where the saccade is launched from. Thus, a model that accounts for these

launch-site dependent differences in exploration biases has the potential to offer a better account of viewing behavior. This stochastic model differs from the random baseline implemented by Najemnik and Geisler (2008), in which they randomly selected each fixation location from all possible points in the display, because it incorporates basic oculomotor heuristics that guide the eyes, without the need for complex computation of peripheral sensitivity or target location probability.

This stochastic search model is related to the more general topic of saccadic biases. Recent work in this area by Le Meur and Coutrot (2016) independently arrived at a very similar model to Clarke et al. (2016) while investigating context-dependent and spatially variant viewing biases. Both their model and the Stochastic Search model partitioned the data into $k \times k$ subsets (Le Meur & Coutrot, 2016, used $k = 3$ while Clarke et al., 2016, used $k = 5$) and then used nonparametric methods to model the distributions. In this paper, we reimplement and generalize this idea with a model named *Saccadic Flow*, and examine the extent to which it is useful as a prior for analyzing eye movements made with more natural (photographic) stimuli over a range of different tasks.

The central bias

There is a strong tendency for people to look close to the center of pictures (Canosa et al., 2003; Clarke & Tatler, 2014; Tatler, 2007; Tatler, Baddeley, & Gilchrist, 2005) and videos (Loschky, Larson, Magliano, & Smith, 2015; Tseng, Carmi, Cameron, Munoz, & Itti, 2009) presented on computer screens. There have been a number of suggestions for why this might be, the simplest being that the center of the stimulus array is the best place to look to make use of parafoveal vision. Another possible explanation for this effect is that the muscles of the eye show a preference for the "straight ahead" position, recentering in the orbit of the eye socket for most comfortable contraction of the ocular muscles—an *orbital reserve* (Fuller, 1996). As most scene viewing experimental set-ups stabilize the head to increase the accuracy of the eye tracking, and most scenes are presented in the center of computer displays, such a recentering mechanism would mean that the center of images would indeed be preferentially selected. However, when scenes are scrambled into four quadrants, fixations are located near to the center of each quadrant, rather than the display center (Stainer et al., 2013), suggesting that the central tendency is responsive to the scene itself rather than to the frame of the monitor upon which the scene is displayed.

Another possible explanation for the central fixation bias is that it represents a response to *photographer bias* in scenes, as photographers tend to frame their shots to include the most important content in the center of the scene. However, when Tatler (2007) presented scenes where the image features were biased towards the edge of the scene, the central fixation bias persisted. The final possibility is that as a consequence of repeated exposure to photographer bias, the center of scenes is simply where people are *trained* to look at images (Parkhurst, Law, & Niebur, 2002). Such learning of spatial probabilities of targets can explain why, for example, people tend to look around the horizon when searching for people in natural scenes (Birmingham, Bischof, & Kingstone, 2009; Ehinger, Hidalgo-Sotelo, Torralba, & Oliva, 2009; Torralba, Oliva, Castelhano, & Henderson, 2006). Expecting to find interesting content in the center of scenes might be a consequence of this hypothesis typically being correct.

Irrespective of why it occurs, Clarke and Tatler (2014) showed that the characteristics of the central bias are remarkably consistent across a series of eye movement databases over tasks such as free-viewing, visual search, and object naming. They proposed a simple, standardized central baseline based on a multivariate Gaussian, and demonstrated that it outperforms similar measures previously used in the literature.

Other behavioral biases in saccades

While the central bias has attracted the most attention (at least in terms of models of visual attention), a number of other biases have been documented. These are discussed below.

Horizontal saccades:

Several researchers have noted that when viewing scenes there is a higher proportion of eye movements in horizontal directions than vertical or oblique movements (e.g., Foulsham, Kingstone, & Underwood, 2008; Gilchrist & Harvey, 2006; Lappe, Pekel, & Hoffmann, 1998; Lee, Badler, & Badler, 2002; Tatler & Vincent, 2008). There are a number of possibilities as to why this tendency exists. Firstly, there may be a muscular or neural dominance making oculomotor movements in the horizontal directions more likely. Secondly, the characteristics of photographic images may mean that content tends to be arranged horizontally by the photographer. In such situations, horizontal saccades may be the most efficient way to inspect scenes. Thirdly, using horizontal saccades in scene viewing might be a learned strategy. Observers may learn the natural characteristics of scenes based on previous experience, and therefore demonstrate an increased likelihood of moving in the horizontal direction. A final explanation is that this tendency is a

consequence of the aspect ratio of visual displays, which normally allow for larger amplitude saccades in the horizontal than in vertical directions (von Wartburg et al., 2007). Results from Foulsham et al. (2008) and Foulsham and Kingstone (2010) suggest that the outline of the displayed scene has a marked effect on saccade directions during viewing. Indeed, Foulsham et al. (2008) found that when the orientation of an image is rotated, the distribution of saccade directions follows the orientation of the scene. Furthermore, when a scene is presented in a circular aperture, the tendency to make horizontal saccades disappears, being replaced by a tendency to make vertical saccades relative to the image orientation (Foulsham & Kingstone, 2010). However, when using fractal images (where images do not have an obvious orientation), observers tend to make horizontal saccades, regardless of the angle that the image is presented. These findings suggest that directional biases in saccades are influenced not only by the shape of the displayed scene but also by its content.

Coarse-to-fine:

Another robust pattern in human saccadic behavior is the tendency to make large eye movements after the initial scene onset, and smaller saccades as the trial unfolds (Antes, 1974; Over, Hooge, Vlaskamp, & Erkelens, 2007; Pannasch, Helmert, Roth, Herbold, & Walter, 2008). This is often accompanied by an increase in fixation durations, and is framed as a move from ambient to focal processing (Follet, Le Meur, & Baccino, 2011; Unema, Pannasch, Joos, & Velichkovsky, 2005; Velichkovsky, Rothert, Kopf, Dornhöfer, & Joos, 2002). Godwin, Reichle, and Menneer (2014) successfully replicated these findings, but they offered an alternative explanation, namely that this behavior is driven by stochastic factors that govern eye movements.

Leftwards bias:

Several studies have shown that observers exhibit a bias to fixate the left half of a stimulus over the right (Brandt, 1945; Learmonth, Gallagher, Gibson, Thut, & Harvey, 2015; Nuthmann & Matthias, 2014; Ossandón, Onat, & König, 2014; Zelinsky, 1996). This effect falls under the more general spatial attention bias of pseudoneglect (Bowers & Heilman, 1980), which also affects tasks such as line bisection. The leftwards bias is typically short-lived, affecting only the first couple of saccades after scene onset, and whereas it is robust, it is comparatively weak compared to other biases in scene viewing. For example, Dickinson and Intraub (2009) found 62% of initial saccades were directed to the left half of the image during free viewing. There is some evidence that this bias is related to native reading direction (Friedrich & Elias, 2014).

Saccadic momentum and inhibition of return:

Several studies have described sequential dependencies during free viewing that bias saccades to repeat the same vector and amplitude (known as saccadic momentum) and to bias saccades away from returning to previously visited targets (known as inhibition of return). Although both of these phenomena bias fixations away from previously fixated locations, they differ in that inhibition of return is bound to a location in the search array; i.e., it is coded in object-based or spatiotopic coordinates (e.g., Krüger & Hunt, 2013), while saccadic momentum has been characterized as a basic tendency to repeat the same motor program (Wang, Satel, Trappenberg, & Klein, 2011). Inhibition of return, unlike saccadic momentum, is task-dependent (Dodd, Van der Stigchel, & Hollingworth, 2009) and is disrupted by removing the scene or inhibited object (Klein & MacInnes, 1999; Takeda & Yagi, 2000). MacInnes, Hunt, Hilchey, and Klein (2014) observed both of these mechanisms operating during free visual search of a complex scene, but presumably only saccadic momentum would be consistently observed for all tasks and images.

The present study

These biases, and in particular the central bias, are important to take into account when evaluating the performance of models of fixation location, and investigating relationships between eye movement data and other factors. The main contribution of this manuscript is to introduce the *saccadic flow* model. This can be thought of as a generalization of the central bias: Instead of simply characterizing the image-independent probability of fixating (x_i, y_i) we model the conditional probabilities $p(x_i, y_i|x_{i-1}, y_{i-1})$, i.e., the probability of making a saccade from to (x_i, y_i) given we are currently fixating (x_{i-1}, y_{i-1}).

In Modeling biases we describe the saccadic flow model and an improved version of the central bias model. The model's ability to account for eye movements during scene viewing is evaluated over 15 previously published datasets. These cover a range of types of images and viewing tasks. In Using biases we demonstrate how the central bias and saccadic flow can be used to improve analysis and visualization methods. In particular, we present bias-weighted gaze landscapes, and demonstrate an interaction between the likelihood of a saccade under different bias models and bottom-up visual salience. Finally, we investigate the

	Observers	Images	Task	Display duration
Clarke et al. (2013)	24	100	Object naming	5000 ms
Yun, Peng, Samaras, Zelinsky, and Berg (2013)—SUN	8	104	Image description	5000 ms
Tatler et al. (2005)	14	48	Memory	Variable
Einhäuser, Spain, and Perona (2008)	8	93	Object naming	3000 ms
Tatler (2007)—free	22	120	Free viewing	5000 ms
Judd et al. (2009)	15	1003	Free viewing	3000 ms
Yun et al. (2013)—PASCAL	3	1000	Free viewing	3000 ms
Tatler (2007)—search	30	120	Visual search	5000 ms
Clarke et al. (2009)	7	360	Visual search	Variable
Ehinger et al. (2009)	14	912	Visual search	Variable
Asher et al. (2013)	25	120	Visual search	Variable
Jiang et al. (2014)	16	500	Free viewing	5000 ms
Borji and Itti (2015)	120	4000	Free viewing	5000 ms
Mills et al. (2011)	53	67	Various	6000 ms
Koehler et al. (2014)	80	800	Various	2000 ms

Table 1. Summary of the 15 datasets used throughout this study. *Note*: The top eight datasets were used to train the model, whereas the bottom seven were used only for evaluation.

shortcomings of these generative models by comparing synthesized data to human eye movements.

Modeling biases

In this section, we (a) update the central bias model of Clarke and Tatler (2014) to make use of a truncated Gaussian distribution that allows us to take the image boundaries into account; (b) explore the strength of the leftwards bias in relation to the central bias; and (c) describe the saccadic flow model.

Modeling methods

Here, we give an overview of the methods and data used for the saccadic flow modeling.

Datasets

We used a number of previously published datasets, covering a range of tasks, images, and experimental set-ups. This variety allows us to produce a model that will generalize well to other datasets. The models were trained on eight of the ten datasets used in Clarke and Tatler (2014). We chose to remove the data from Asher, Tolhurst, Troscianko, and Gilchrist (2013) from our training set as the images have an aspect ratio of 5:4, whereas the rest of the data in our training set has an aspect ratio of 4:3. The pedestrian search dataset (Ehinger et al., 2009) was removed from the training set as previous analysis (Clarke & Tatler, 2014) shows that it is biased compared to the other datasets analyzed.

Both of these datasets were used as test sets to evaluate how well our models generalize.

We also added four new datasets to the ten used by Clarke and Tatler (2014). These were used to test the model.

- Jiang, Xu, and Zhao (2014) collected data from 16 observers viewing 500 natural scenes containing crowds of people (aspect ratio 4:3).
- Clarke, Chantler, and Green (2009) investigated visual search for a target on a homogeneous textured background (i.e., target in noise). This dataset differs from the previous in that there is no semantic image content in the scene, and the stimuli had a 1:1 aspect ratio.
- Greene, Liu, and Wolfe (2012) released a dataset of observers viewing square grayscale photographs.
- Borji and Itti (2015) recently released a very large ($\approx$ 650,000 fixations, 2,000 images) dataset collected over twenty different stimulus types. Given the size of this dataset, and the wide-screen 16:9 aspect ratio, the evaluations on this dataset are presented separately, and split by stimulus class.

This gives us a relatively homogeneous training set, and a more heterogeneous test set. Hence, good performance on the test sets will likely be indicative of a generalizable result. An overview of the datasets used is given in Tables 1 and 2.

Preprocessing

As with Clarke and Tatler (2014), we normalized all fixations to the image frame, keeping the aspect ratio constant: that is, $(x, y) \in (-1.-1) \times (-a, a)$ with typically $a = 0.75$. The initial saccades after image

	Eye tracker	Viewing distance	Screen size	Image size	Viewing angle	Chin rest
Clarke et al (2013)	EyeLink II	50 cm	21"	800 × 600	31 × 25 deg	no
Yun et al (2013)- PASCAL	EyeLink 1000	?	?	?	?	?
Tattler et al (2005)	EyeLink I	60 cm	17"	800 × 600	30 × 22 deg	no
Einhauser et al (2008)	EyeLink 1000	80 cm	20"	1024 × 768	29 × 22 deg	yes
Tattler et al (2007)- free	EyeLink II	60 cm	21"	1600 × 1200	40 × 30 deg	no
Judd et al (2009)	?	2 feet	19"	1024 × 768	?	yes
Yun et al (2013)- SUN	EyeLink 1000	?	?	??	?	?
Tattler et al (2007)- search	EyeLink II	60 cm	21"	1600 × 1200	40 × 30 deg	no
Clarke et al (2009)	Tobii x50	87 cm	20"	1024 × 1024	16.7 × 16.7 deg	yes
Ehinger et al (2009)	ISCAN RK-464	75 cm	21"	800 × 600	23.5 × 17.7 deg	yes
Asher et al (2013)	& EyeLink 1000	55 cm	?	1024 × 1280	37.6 × 30.5 deg	yes
Jiang et al (2014)	& Eyelink 1000	57 cm	22"	1024 × 768	38.8 × 29.1 deg	?
Borji and Itti (2015)	Eyelink 1000	106 cm	42"	1920 × 1080	45.5 × 31 deg	yes
Mills et al (2011)	EyeLink II	44 cm	?	?	16.9 × 22.8 deg	no
Koehler et al (2014)	Eyelink 1000	40 cm	?	800 × 600	15 × 15 deg	no

Table 2. Details of the experimental setups in each of the 15 datasets analysed in the present study. *Notes*: We provide only information reported in the original articles. Question marks indicate information not reported in the original article. *For the Judd et al. dataset, images varied in pixel dimensions but the majority were at 1024 × 768.

onset (9.1% of the data) were excluded, giving us a total of 159,226 saccades. Saccades with a start or end point falling outside of the image frame were also removed.

When fitting saccadic flow models, we *mirrored* the set of fixations, by adding in horizontally and vertically reflected copies of the data. This has two advantages: First, it is an easy way to make the saccadic flow bias symmetric in the horizontal or vertical directions. This is similar to how the central bias was defined by Clarke and Tatler (2014). Second, it increases the amount of data available for fitting by a factor of four. This is important as (due to the central bias) there are relatively few saccades that originate from the corners of the images. By equating all corners, we can pool the data and obtain more stable estimates for the underlying distribution. The downside of mirroring saccades in this manner is that our model of saccadic flow will be insensitive to the *leftwards bias* in natural scene viewing (Nuthmann & Matthias, 2014). However, as this accounts for a relatively small proportion of the overall variance in the data (Left versus right), we view this as an acceptable tradeoff. Similarly, as we do not factor in the timecourse of the scan-path, we will not capture *coarse-to-fine* dynamics (saccadic amplitude tends to decrease with time from stimulus onset).

Truncated central bias

As the first step in modeling saccadic flow, we will update the central bias from Clarke and Tatler (2014) and use a truncated normal distribution. This is

straightforward. Refitting a multivariate Gaussian to the data reduces the deviance in the central bias model by 4.4%. Using a truncated Gaussian gives us an improvement of 12%. We can round the truncated Gaussian model to $\mu = (0, 0)$, with a covariance matrix of (0.32, 0; 0, 0.144) with no loss of precision. That is, this is identical to Clarke and Tatler (2014) except with $\sigma = 0.32$ rather than 0.22. We will use the abbreviations CT2014 and CT2017 to refer to these models.

Left versus right

As mentioned above, the downside of mirroring the saccades in our dataset is that our bias model will be symmetric and will be unable to exhibit the leftward bias observed in human fixation data. Here, we investigate the size of the leftwards bias (in the unmirrored data) by plotting how the distribution of horizontal fixation location varies with fixation number (Figure 2). We can see that while we do have a leftwards bias in our data, it is a small effect that only lasts for the first five fixations after scene onset. Furthermore, there is no sign of an asymmetry in the vertical direction. Fitting an ANOVA to predict the x coordinates of the fixations, given the fixation number, produces adjusted $R^2 = 0.004$. If we limit our analysis to the first five fixations in each scan-path, this only increases to adjusted $R^2 = 0.01$. The small size of the R^2 in both instances suggests that by ignoring the leftwards bias, we lose little explanatory power. This brings the advantage of then allowing us to treat everything as symmetrical, which simplifies the model
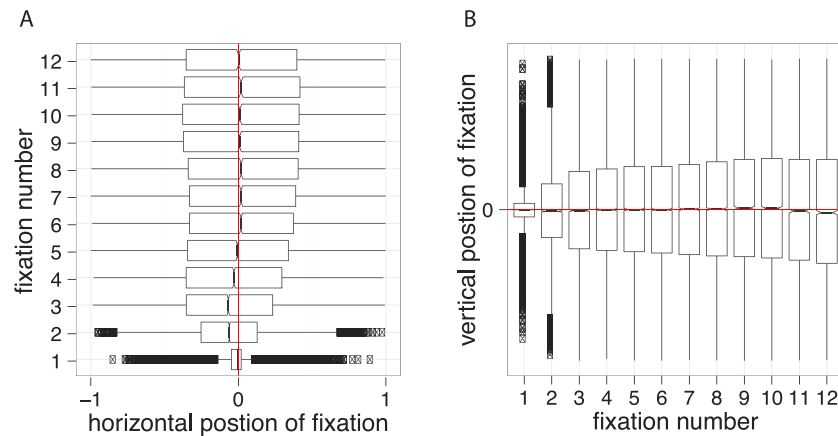
Figure 2. Boxplots showing the distribution of horizontal and vertical fixations by fixation number in the merged training set.

and increases the amount of data available (by mirroring fixations).

Saccadic flow

Saccadic flow can be thought of as a generalization of the central bias, and is illustrated in Figure 1. Instead of computing the distribution of all saccadic endpoints in a dataset, we look at the distribution of saccade endpoints given the start points, i.e., for a saccade from (x_0, y_0) to (x_1, y_1) we want to model $p(x_1, y_1 | x_0, y_0)$.

Modeling

To characterize how the distribution of saccadic endpoints varies with the start point, we used a sliding window approach. All saccades that originated from an $n \times n$ window were taken and used to fit a truncated multivariate Gaussian distribution using the tmvtnorm library for R. This window was moved in

steps of $s = 0.01$ from $[-1, -0.75]$ to $[1-n, a-n]$. Windows containing less than 250 datapoints were discarded. We experimented with varying the window size ($n \in \{0.05, 0.1, 0.2\}$). However, as this parameter was found to have a negligible result, we only report the results for $n = 0.05$.

Multivariate polynomial regression was then used to fit fourth-order polynomials to each of the parameters. As polynomial regression performs poorly in the presence of outliers, we will also use robust estimation (rlm from the MASS library). This will stop the model fits being overly influenced by outlier points from the image boundary. Figure 3 shows how the parameters for the truncated multivariate Gaussian distributions vary over horizontal position for a selection of vertical positions. The regression coefficients (given in supplementary materials) allow us to estimate the conditional probability of a saccade to (x_1, y_1) given the starting fixation (x_0, y_0). As the robust estimation methods give a far better fit to the data, we will use this version of the model and discard the polynomial regression version.
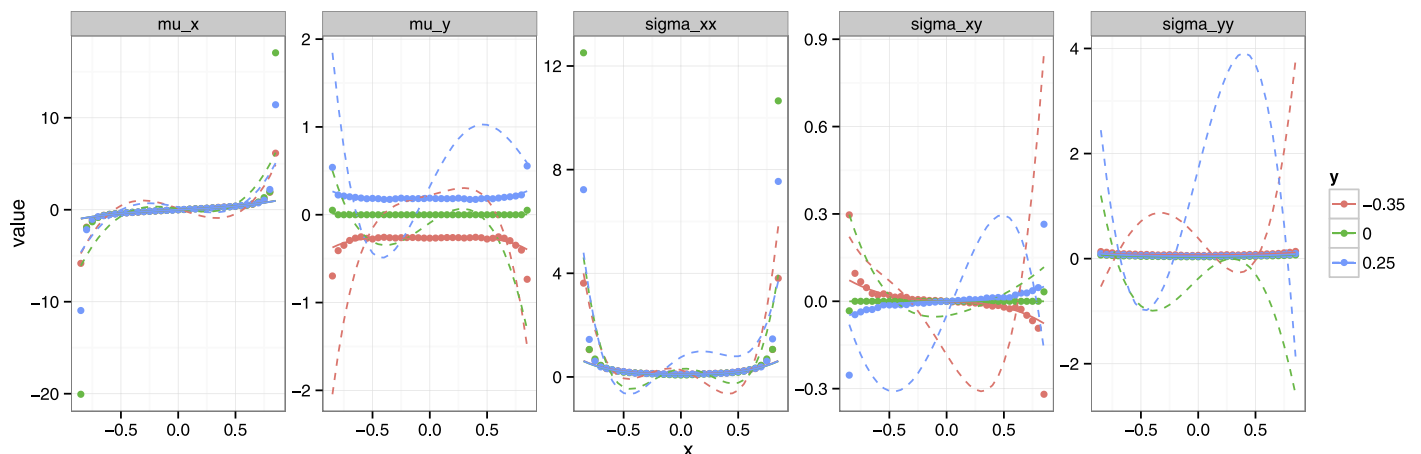

Figure 3. How the truncated Gaussian parameters vary with saccadic starting location. Dotted line shows polynomial regression fits; solid line shows robust polynomial regression.

Evaluation will be done using bootstrapping (100 repetitions with $N = 1000$). Not only does this allow for confidence intervals to be estimated for our result, but it also allows us to sidestep the problems of using datasets of very different sizes: Likelihood scores are heavily influenced by the number of points included in the analysis, so having this fixed at $N = 1000$ means we can compare across datasets more meaningfully. We chose to evaluate how well the various models work by simply calculating the likelihood, $p(\text{data}|\text{model})$, for each dataset, and reporting the difference in log likelihood between a uniform distribution and our models. As the number of datapoints is much larger than the number of parameters, log likelihood approximates AIC. We also report a receiver operator curve, ROC (Green & Swets, 1966), analysis in which we look at how often saccades land within the most likely $x\%$ of the prediction maps from the different bias models. This analysis was done on the complete datasets without bootstrapping.

Results

How well does this model account for the fixations in our datasets? Figures 4 and 11 compare how well the different models outperform a uniform distribution in terms of log-likelihood. We can see that in all cases, the flow model offers a much larger improvement over a uniform distribution than either central bias model. The differences between the two central biases is much smaller, but in general, we can see that using a truncated distribution (to correctly take the image boundaries into account) offers a small improvement over the Clarke and Tatler (2014) bias.

It is interesting to note that the flow model still does a good job of accounting for the distribution of saccades in datasets (those involving visual search) in which the central bias is outperformed (in terms of log-likelihood) by the uniform distribution, chiefly the data from Asher et al. (2013), Clarke et al. (2009), Tatler (2007). We can see a very similar pattern of results in the ROC analysis in Figure 5.

The effect of task

We now examine how the ability of saccadic flow to explain different scan-paths depends on the observers' task. We will make use of datasets from Mills, Hollingworth, Van der Stigchel, Hoffman, and Dodd (2011) and Koehler, Guo, Zhang, and Eckstein (2014). To look at task effects we computed the mean log likelihood over each scan-path in these datasets (see Figure 6). We can see that while task has a slight influence over the mean log likelihood for a scan-path, there is a large degree of overlap in the distributions. Additionally, we can see that the relative likelihood of
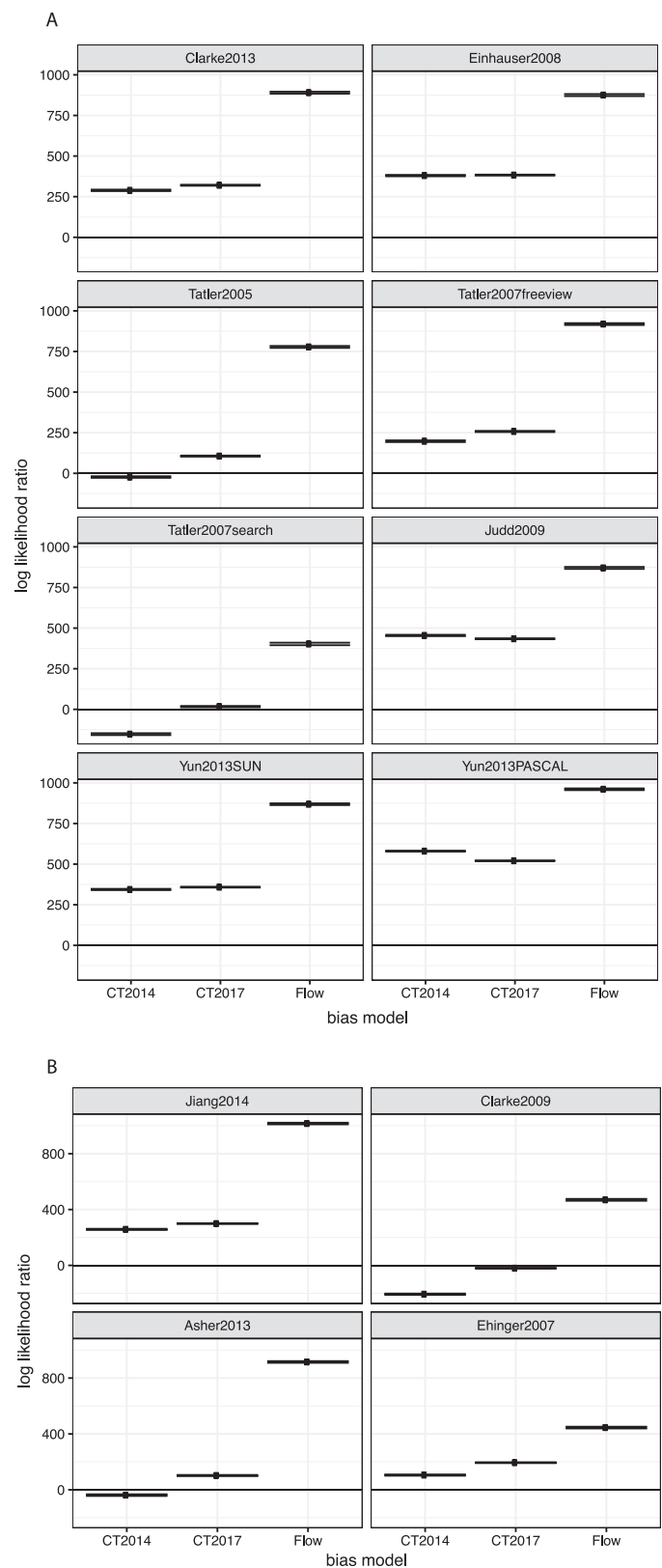

Figure 4. Modeling results. We can see that the flow model offers a much larger improvement in terms of log-likelihood than either of the central bias models. This holds even in datasets which do not show a strong central bias.
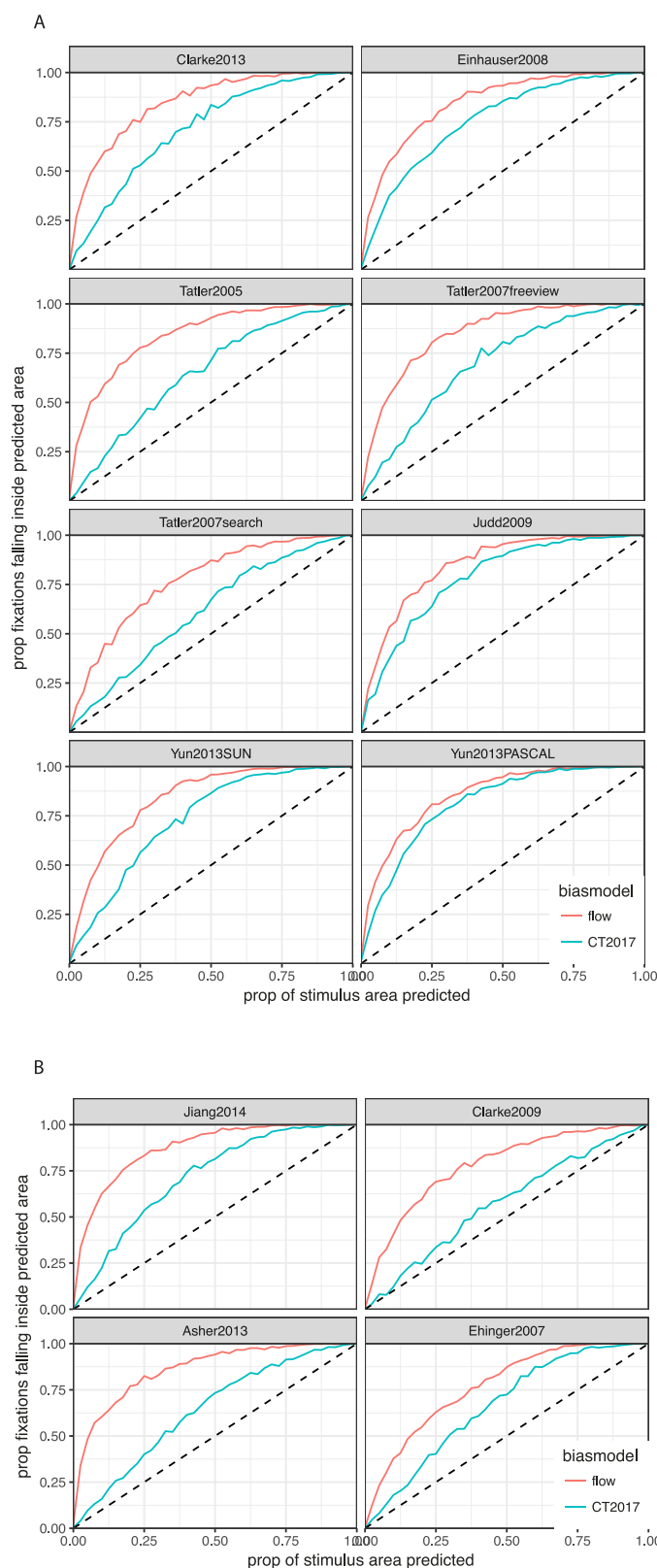
Figure 5. ROC analysis comparing the flow model to the central bias.

scan-paths made during visual search compared to other tasks varies between datasets. These results suggest that at least for the datasets considered here, the extent to which saccadic flow is able to explain the observed scan-paths is not strongly influenced by the observer's task.

Saccadic flow and underlying physiology

The saccadic flow model that we develop and evaluate here is a statistical model developed based on fitting empirical data. As such it does not make explicit any underlying physiological constraints or neuro-physiological architecture. However, by constructing the model from observed data, these underlying constraints are necessarily present in the statistical model that we present here. The anisotropies in saccade directions that underlie the construction of the saccadic flow model are likely to reflect a combination of responses to the image and physiological constraints imposed by the arrangement and action of oculomotor muscles (Smit, Van Gisbergen, & Cools, 1987; Viviani, Berthoz, & Tracey, 1977). Similarly, the skew in saccade amplitudes toward favoring small amplitude saccades may reflect aspects of the drop off in acuity limits with distance into the peripheral retina. Thus while these biomechanics and neurophysiological factors are not explicit in saccadic flow, they necessarily inform the construction and thus any predictions arising from the model.

Using biases

This section makes use of an improved central bias model and the *saccadic flow* model (described in Saccadic flow). The new central bias model is similar to the model presented by Clarke and Tatler (2014), except for using a truncated Gaussian distribution to take the image boundaries into account. We present three examples of how these bias models can be used as priors in order to weight fixations, based on the fact that Flow produces likelihoods for any given fixation given the current fixation. First of all, we will demonstrate how we can weight fixations in gaze landscapes (also known as hotspot maps or heatmaps) to reduce noise and to give an improved visualization of the image regions participants looked at more than expected. Secondly, we examine whether saccadic flow can be used to better understand the contribution of low-level features on fixation selection, and potentially lead to better evaluation of such computational saliency models. Finally, we demonstrate how flow can be used to generate a series of saccades, and compare these to observed human saccades. Being able to
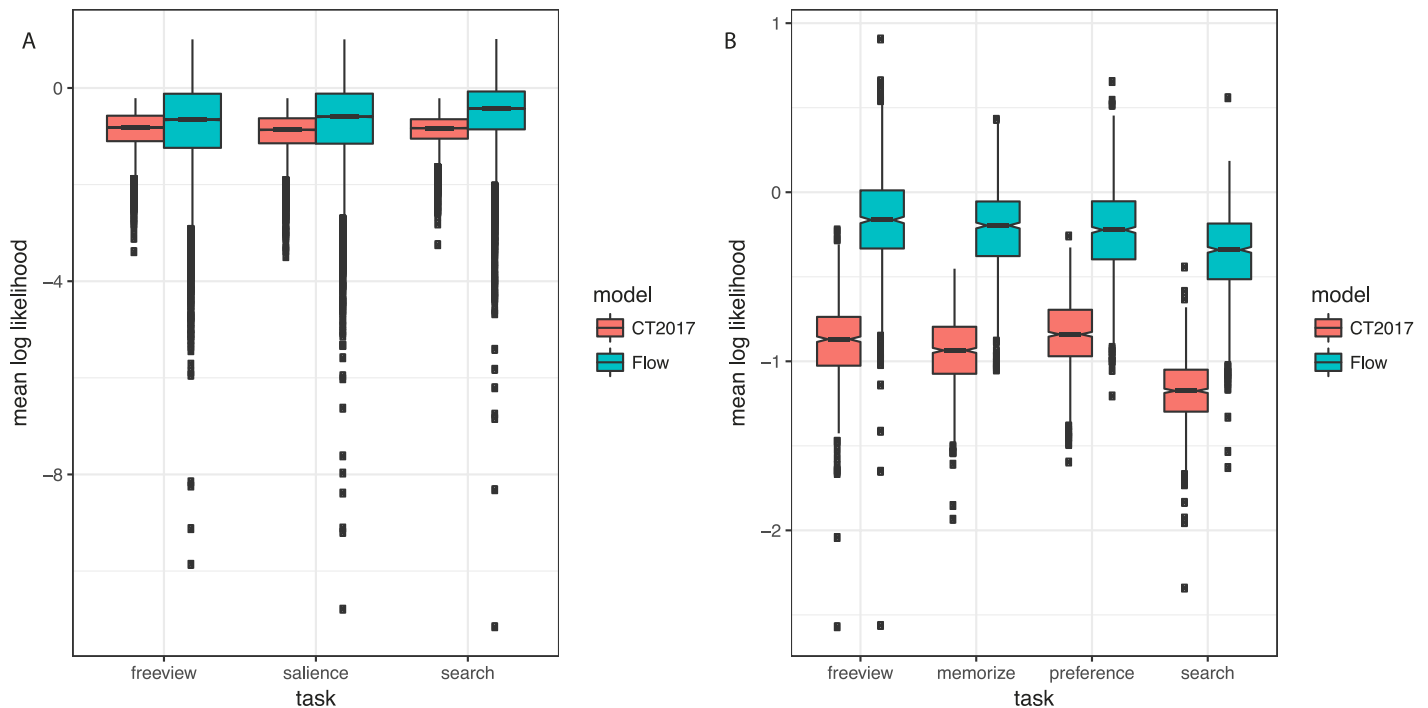
Figure 6. The influence of task on the extent to which saccadic flow can explain scan-paths for the (a) Koehler et al. (2014) and (b) Mills et al. (2011) datasets.

generate realistic synthetic datasets is useful to create an image-independent baseline with which to examine spatial maps of prediction using signal detection theory (see Clarke & Tatler, 2014).

Gaze landscapes

One technique that is commonly used to visualize the spatial allocation of gaze is to create "heatmap" plots where color or luminance is used to indicate the density of fixation on those locations (Figure 7, column 2). A potential problem with visualizing data in this way is that such maps represent all fixations as being of equal importance. For example, a location that is fixated for one second would be weighted equally with fixations that lasted half that time. If we want to make an assumption that fixation duration is intimately linked with the importance of that fixation (i.e., we will look longer at more informative information), then we can change our visualization to weight fixations by their duration (Figure 7, column 3). However, this weighted heatmap still fails to distinguish fixation behavior likely to arise from image independent biases like the central fixation bias from fixation behavior likely to reflect meaningful interrogation of, and response to, the viewed content.

An advantage of the Clarke and Tatler (2014) model and the saccadic flow model presented here is that we can represent fixations by the likelihood that they would occur based on the predictions of the models.

Because the models reflect image-independent behavioral and oculomotor biases, fixations not predicted by these models might involve more high-level mechanisms. For example, given a tendency to fixate in the center of the scene, we might consider saccades to noncentral locations to be less predicted and therefore more likely to be image- or task-related. In Figure 7 (column 4 and 5) we present some overlaid heatmap data from the Clarke et al. (2013) dataset, where fixations are weighted by the inverse probability of them occurring based on the models of central bias and saccadic flow. These figures reveal that representing data in this manner can allow us to visualize information that was important enough to disrupt these biases. In other words, these visualizations remove some of the image-independent biases, and reveal the more important image *dependent* information.

The top row of Figure 7 demonstrates that weighting the fixations by the central bias and flow model both reduce the *importance* of some fixations. The central bias model punishes fixations near the center of the image, while the flow model punishes fixations that were well predicted by the oculomotor biases of the saccadic flow model. Conversely, the models reward unlikely fixations. The second row reveals an instance of where the car to the left received fewer fixations than the pub sign, but that these fixations are boosted in the central bias and saccadic flow models where "unlikely" saccades were made to this location. In the third and fourth rows, there are examples of images with important content near the center of the photograph.
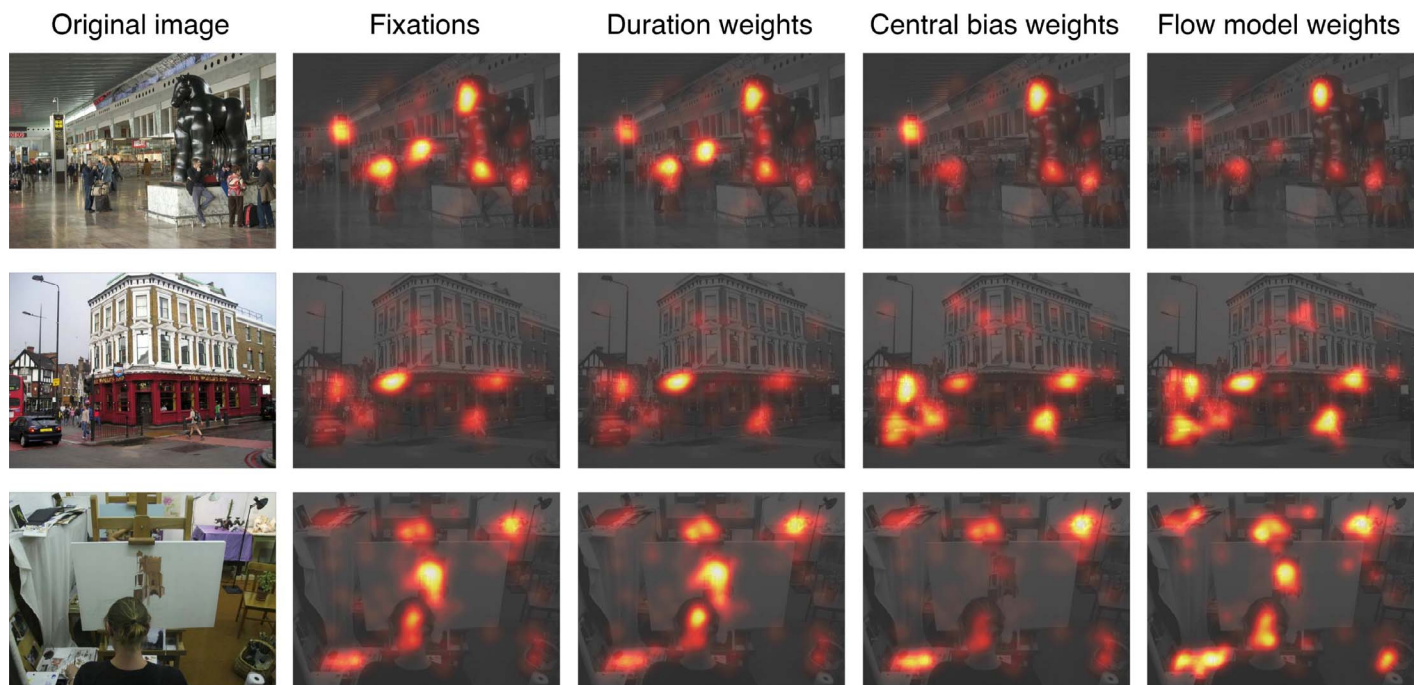
Figure 7. Examples of fixation heatmap plots from Clarke et al. (2013). The same fixations are presented where the Gaussian at each fixation is weighted by the duration of the fixation, the center bias model from Clarke and Tatler (2014), and the saccadic flow model presented in this paper.

This illustrates how the central bias model can sometimes overcompensate and reduce the influence of fixations in the center of pictures that have important content located there. Given the tendency for photographers to center their photographs around important content, reducing the weight of fixations to the castle in the painting (row 3) and the girl's face (row 4) would perhaps overly punish centrally biased photographic composition. With the flow model, however, these areas are still represented, as observers made saccades to these regions that were unlikely to be driven by behavioral biases.

Removing biases when examining image-dependent information

By considering saccades in light of the probability that they were generated by image-independent biases, we can gain further insights into the image-dependent features that are important in attracting fixation. One feature that has been shown to correlate with fixation is visual salience (Parkhurst et al., 2002). However, others have argued that this tendency is driven by the correlation between salient objects and their semantic interest (Henderson, Castelhano, Brockmole, & Mack, 2007), with interesting objects tending to be placed near to the center of photographs (Tatler, 2007). Oculomotor biases which favor a central tendency would predict the same fixation placement regardless of the distribu-

tion of salient objects in the image (Tatler & Vincent, 2009). Here, we can examine this question by looking at the relationship between saccade probability and the ability of different conspicuity maps to predict fixation. We can therefore examine how the effect of visual salience observed in eye movement analysis is related to the behavioral biases of eye movements.

We compared the proportion of fixations that fell in the brightest 20% of pixels for salience maps to the likelihood of fixations from the flow and central bias models. Fixations were separated for each image into bins of 5% from the least likely to the most likely to be generated based on salience. We then examined what proportion of each of these bins were in the brightest 20% of salience maps using the Adaptive Whitening Saliency (AWS; Garcia-Diaz, Leborán, Fdez-Vidal, & Pardo, 2012), RARE (Riche et al., 2013) and Graph-based visual saliency (GBVS; Harel, Koch, & Perona, 2006) algorithms. We selected AWS and RARE as they are the two best performing salience models according to the MIT Saliency Benchmark (Bylinskii et al.; Judd, Durand, & Torralba, 2012) with publicly available code, and GBVS as it contains a bias towards the center caused by summing neighboring pixel values across the spatial prediction map. Figure 8 reveals that the likelihood of making a saccade based on both the central bias and the flow model is highly related to salience in both AWS and RARE, with low-likelihood saccades being less likely to be to a salient region. Saccades that are very unlikely to be generated based
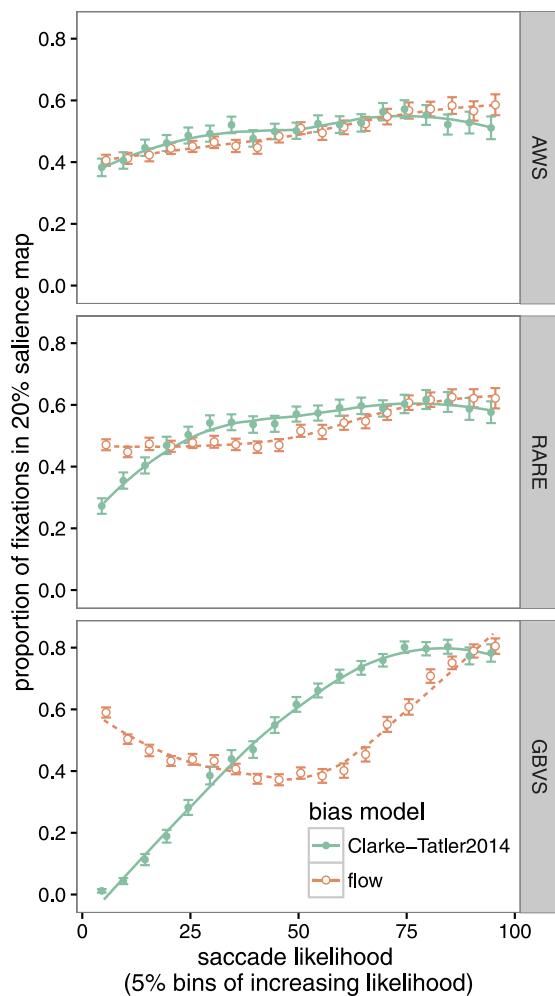
Figure 8. Saccades binned by probability of them occurring in 5% bins against the proportion of those fixations that fell in a 20% thresholded region of AWS, RARE and GBVS salience maps.

on the oculomotor tendencies of eye movement (both flow and central bias) are therefore also less well explained by salience. Of the 5% of fixations that were *most* likely from saccadic flow, 60% of fixations fell in the 20% thresholded region of the AWS map. However, of the 5% of fixations that were *least* likely from saccadic flow, only 40% of fixations fell in this region. This means that it may be important to consider, and potentially remove, behavioral biases when attempting to predict fixation selection using feature-based models to ensure that any benefit in predictive power cannot be

explained by behavioral biases correlating with salience. When examining a model that contains an inherent central bias (GBVS), we can see that weighting fixations by the Clarke and Tatler (2014) central bias model is highly related to the performance of GBVS in predicting fixation selection.

Saccadic flow as a generative model

Another use of the saccadic flow model is that it allows us to make spatial maps that relate to the probability of *all saccades within a scene* based on the current position. For example, Figure 9 shows that for three fixations in different locations within a scene, flow will make different spatial predictions of the next saccadic landing position. We can use this method to generate sequences of synthetic scan-paths. Here, we compare the distributions of these generated scan-paths with empirical scan-paths to determine which aspects of human saccadic behavior are not captured by our model. To do this, we will create a merged dataset of fixations from the eight training datasets (175,000 fixations, including initial fixations, in total over 16,000 trials), and then generate a matched synthetic dataset such that the number of fixations in each trial is identical.

We can see from Figure 10a and b that both the central bias and the saccadic flow model do a good job of capturing the distribution of fixation locations over the *x* and *y* axes. While it is not surprising that the central bias closely matches the empirical distributions (as this is exactly what it has been fitted to), it is interesting that saccadic flow does just as good a job. Hence, the central bias can be thought of as a property of saccadic flow, and does not need to be accounted for separately.

When compared to the empirical distributions, both the central bias and saccadic flow appear to be slightly biased towards making fixations to the extreme edges of the image. This suggests that the truncated Gaussian distribution does not quite capture the effects of the image boundary on fixation selection and there is some additional aversion to fixating close to the screen edge.

Another discrepancy between the synthetic and empirical distributions can be seen with saccadic amplitudes. While the flow model is a better fit to the human data than the central bias, it still underestimates the proportion of very short saccades (Figure 10c).
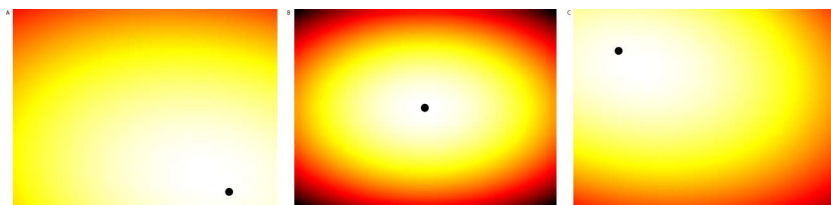

Figure 9. Example spatial prediction maps for all potential saccade locations from three different fixation positions (black circles) to demonstrate how flow's predictions differ across the extent of a scene.
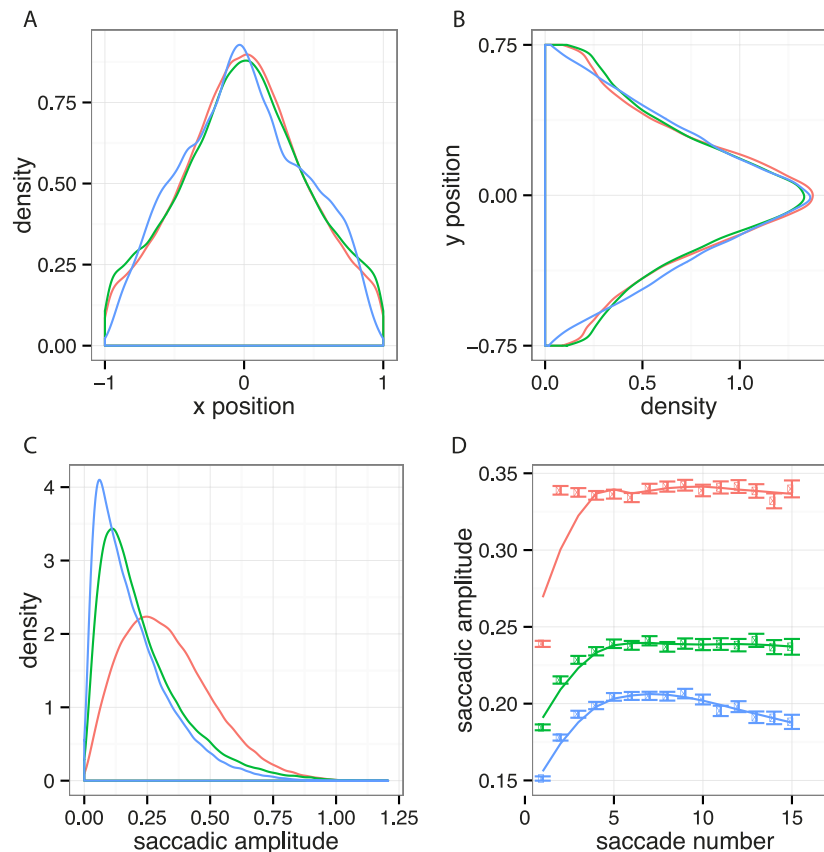
Figure 10. *Blue*: human; *red*: central bias; and *green*: saccadic flow. *Top row*: Comparison of *x* and *y* fixation positions between human fixations and synthetic points generated from the central bias and flow model. *Bottom row*: We can see that the flow model consistently makes saccades with a slightly larger amplitude than do human observers. Distances are expressed relative to the width of the image. Best fit line in (d) fitted with loess regression. All distances are given in normalized units in which the width of an image is 2 (see Modeling methods).

Interestingly, the flow model does manage to capture the initial increase in saccadic amplitudes after scene onset (Figure 10d), but it does not explain the subsequent coarse-to-fine dynamics that are seen in the empirical scan-paths.

Summary

We have demonstrated different ways biases such as saccadic flow and the central bias can be used in eye movement research. They can be used as a prior on the probability of making saccades to different regions of the image, allowing us to then more clearly visualize the image-dependent behavior. We have also shown that the likelihoods of fixations under the bias models are related to features such as salience. The interpretation of visual salience as a predictive model of fixation selection can therefore be informed by considering how likely a saccade is to be generated by these models. Finally, we can also use the bias distributions to generate synthetic data that can be used as control points in ROC analysis, and to explore which aspects of

human saccadic dynamics are not captured by the simple flow model.

Discussion

There has been much effort to generate a predictive model of human eye movements (Bylinskii et al., Judd, Ehinger, Durand, & Torralba, 2009). We propose the saccadic flow model as a robust prior for the image-independent saccadic behavior that is evident when people look at pictures (Tatler & Vincent, 2009). Our Saccadic Flow model provides a better account of eye movement behavior across 15 published datasets than the original Clarke and Tatler (2014) Central Bias model, and a new version of this model using truncated multivariate normal distributions. We find that Saccadic Flow accounts for eye movements across many different tasks and image types. As saccade probabilities across a scene can be predicted by Saccadic Flow, we therefore provide a method of modeling oculomotor biases that can be included in combined models of eye
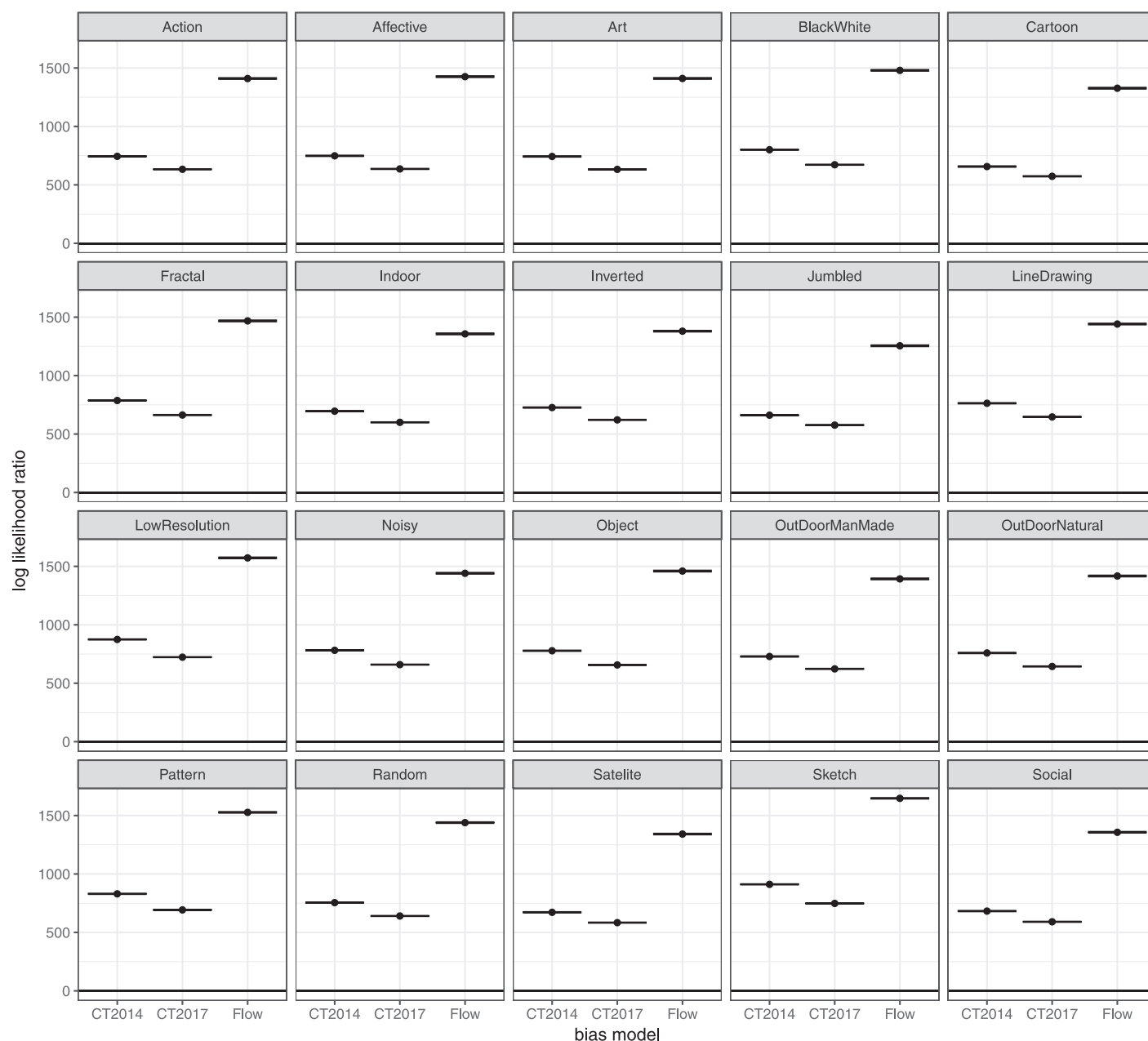
Figure 11. Modeling results for the Borji and Itti (2015) data. We can see that the flow model offers a much larger improvement in terms of log-likelihood than either of the central bias models.

guidance, much as Torralba et al. (2006) and Ehinger et al. (2009) use context maps to predict where people will search for people.

Using saccadic flow

There are two ways in which models of eye movements may benefit from including such information. First, models may include saccadic flow in their calculation of spatial prediction. Understanding where someone is currently fixating in an image appears to dramatically influence where they will go next; as this saccadic flow can be parametrically estimated from any point on an image, it can be used to weight models of low-level (i.e., visual conspicuity) and high-level (i.e., semantic interest) features. Thus, whether someone fixates one of two equally conspicuous, equally interesting objects may be simply determined by the way that the eyes *tend* to move.

The second potential utility of saccadic flow is to generate realistic control fixations with which to evaluate observed fixation data. In this way, saccadic flow can be thought of as a partner to the Clarke and Tatler (2014) central bias, and we expect that in some cases, the simpler central bias will be sufficient (for

example, when examining the overall distribution of fixations rather than the sequence of saccades). However we have demonstrated that while the flow model requires more parameters—we use 16 coefficients to track how each of the five truncated Gaussian parameters vary as a function of (x, y), although many of them are ≈ 0—it generalizes well from one dataset to another and is a far better baseline for modeling a scanpath than the central bias.

In the present paper, we have provided illustrative examples of how saccadic flow can be used to improve our understanding of eye guidance in scene viewing. Heatmaps that account for fixation likelihood under the behavioral biases captured in saccadic flow better reflect the image-dependent biases. Using these to base subsequent analysis of scene content at these locations or of differences in fixation behavior under different tasks allows the researcher to focus analytical efforts on the viewing behavior that is unlikely to arise from image-independent biases in how observers move their eyes. Given the prominence of image-independent biases in observed eye movement behavior (Tatler & Vincent, 2009), removing these biases appropriately from analyses is important for effective evaluation of changes in behavior arising from viewed content or behavioral task. Similarly, any attempt to model the involvement of factors such as image salience in eye guidance, should remove image-independent biases from modeling efforts in order to appropriately evaluate the role of any features under investigation. At present, the utility of removing the central bias from such modeling efforts is widely recognized (Borji & Itti, 2013; Tatler, Hayhoe, Land, & Ballard, 2011), but we have shown in the present work that saccadic flow offers an even better explanation of underlying biases in fixation behavior. Using saccadic flow to remove image-independent biases from datasets of eye movements will be an important improvement for testing existing models of fixation selection, and for better developing new models. Thus while the present work does not provide a direct answer to what factors govern scene inspection, it provides a vital tool for the field to allow this question to be addressed more effectively and appropriately than is currently possible.

Comparisons to existing models

Tatler and Vincent (2009) have previously demonstrated that representing saccade probabilities based on the oculomotor biases in eye movements can account for human fixation behavior during free-viewing reasonably well. Here we extend this concept to spatially adapt the prediction of saccade depending on where in a scene the preceding fixation lies. This is an important step, as it aligns the concept of behavioral

biases in eye movement with the central bias (Tatler, 2007), whereby saccades are likely to be directed towards the center of the image. An advantage over the simpler central bias of Clarke and Tatler (2014) is that Saccadic Flow does not suppress fixations in the center of the screen (where interesting content tends to lie), and sets of saccades generated from the Saccadic Flow model will produce the same centrally biased tendencies as observed fixations.

The work presented here improves on recent models by Clarke et al. (2016) and Le Meur and Coutrot (2016) by offering a parametric model that avoids coarsely partitioning the data into large bins. We have demonstrated that the Saccadic Flow model generalizes well to unseen test datasets, although this is likely to hold only for stimuli that are broadly similar to the images used to train the model (photographs of natural and manmade scenes). As we move away from photographic images to stimuli such as computer interfaces, we would expect the Flow model to offer a poor account of the data (Le Meur & Coutrot, 2016). While we have shown that the model performs well over a small range of tasks (free-viewing, scene description, object naming, and visual search), we do observe differences in the log-likelihood when different tasks are carried out while viewing the same images.

Limitations of saccadic flow

There are several limitations to our modeling work. Firstly, by using a truncated Gaussian, we are unable to capture the skewed nature of the distribution of saccades originating from the corners (see Figure 1). We experimented with fitting a skew-normal distribution using the sn package for R, but met with limited success due to having to deal with the image boundaries. We expect image boundaries is one of the reasons why our saccadic flow model generates saccades with, on average, greater amplitudes than those seen in empirical distributions. The second simplification we make is to not take the leftwards biases in saccadic behavior into account. However, our results suggest that this factor has a relatively small effect on saccades, and is important only for the first couple of saccades. Similarly, our model does not take into account coarse-to-fine biases. Across scene viewing, human fixations tend to increase in duration, and saccade amplitudes tend to decrease as the observer's understanding of the image changes (Antes, 1974). Finally our model only considers the immediately preceding fixation as having an effect. This is likely to be an oversimplification of saccadic programming, given evidence in the literature that previous saccades and fixations influence saccade generation via processes such as saccadic momentum or inhibition of return

(MacInnes et al., 2014). With sufficient data, the modeling framework here could be extended to take the previous *n* fixations into account. It should be noted, however, that the likely impact of these theoretical concepts upon the flow model is unclear: Indeed failing to account for saccadic history may not be important for modeling some aspects of human search behavior (Clarke et al., 2016). The current implementation of saccadic flow also offers an opportunity to empirically assess the likely contribution of such factors as it offers a means to assess the likelihood of repeating a saccade (saccadic momentum) or returning to a previous location (inhibition of return).

Conclusions

Behavioral biases in eye movement are prevalent during scene viewing. Our saccadic flow model allows calculation of saccade likelihood across an image based on empirical data of how the eye tends to move in many different scene viewing conditions, with flow providing a strong fit to several datasets. There are a number of ways that flow can be developed, and we propose that gaining a better understanding of the saccadic biases underlying fixation behavior can only be a positive for our search to understand why people look where they look. Whereas the central bias model may be a better choice in some contexts (i.e., when the analysis is in terms of unordered fixation coordinates), we recommend using saccadic flow where possible. Flow consistently explains more variance than uniform and CT2014/17 models while also accounting for the central bias. This suggests that our model is robust, generalizable, and should be of use to researchers interested in eye movements in a variety of scene-viewing paradigms.

Keywords: eye movements, central tendency, salience

Acknowledgments

Email: a.clarke@essex.ac.uk.
Address: Department of Psychology, University of Essex, Colchester, UK.

References

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*(1), 62–70.

Asher, M. F., Tolhurst, D. J., Troscianko, T., & Gilchrist, I. D. (2013). Regional effects of clutter on human target detection performance. *Journal of Vision*, *13*(5), 25, 1–15, doi:10.1167/13.5.25. [PubMed] [Article]

Birmingham, E., Bischof, W. F., & Kingstone, A. A. (2009). Get real! Resolving the debate about equivalent social stimuli. *Visual Cognition*, *17*(6–7), 904–924.

Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 185–207.

Borji, A. & Itti, L. (2015). Cat2000: A large scale fixation dataset for boosting saliency research. *arXiv preprint arXiv:1505.03581*.

Bowers, D., & Heilman, K. M. (1980). Pseudoneglect: Effects of hemispace on a tactile line bisection task. *Neuropsychologia*, *18*(4), 491–498.

Brandt, H. F. (1945). *The psychology of seeing*. New York: The Philosophical Library.

Buswell, G. T. (1935). *How people look at pictures: A study of the psychology and perception in art*. Chicago, IL: University of Chicago Press.

Bylinskii, Z., Judd, T., Borji, A., Itti, L., Durand, F., Oliva, A., & Torralba, A. MIT saliency benchmark. Retrieved from http://saliency.mit.edu/.

Canosa, R. L., Pelz, J. B., Mennie, N. R., & Peak, J. (2003). High-level aspects of oculomotor control during viewing of natural-task images. *Proceedings of SPIE Human Vision and Electronic Imaging VIII*, *5007*, 240–251.

Clarke, A. D. F., Chantler, M. J., & Green, P. R. (2009). Modeling visual search on a rough surface. *Journal of Vision*, *9*(4):11, 1–12, doi:10.1167/9.4.11. [PubMed] [Article]

Clarke, A. D. F., Coco, M. I., & Keller, F. (2013). The impact of attentional, linguistic, and visual features during object naming. *Frontiers in Psychology*, *4*, 927.

Clarke, A. D. F., Green, P. R., Chantler, M. J., & Hunt, A. R. (2016). Human search for a target on a textured background is consistent with a stochastic

model. *Journal of Vision*, *16*(7):4, 1–16, doi:10.1167/16.7.4. [PubMed] [Article]

Clarke, A. D. F., & Hunt, A. R. (2015). Failure of intuition when presented with a choice between investing in a single goal or splitting resources between two goals. *Psychological Science*, *27*(1), 64–74.

Clarke, A. D. F., & Tatler, B. W. (2014). Deriving an appropriate baseline for describing fixation behaviour. *Vision Research*, *102*, 41–51.

Dickinson, C. A., & Intraub, H. (2009). Spatial asymmetries in viewing and remembering scenes: Consequences of an attentional bias? *Attention, Perception, & Psychophysics*, *71*(6), 1251–1262.

Dodd, M. D., Van der Stigchel, S., & Hollingworth, A. (2009). Novelty is not always the best policy inhibition of return and facilitation of return as a function of visual task. *Psychological Science*, *20*(3), 333–339.

Ehinger, K. A., Hidalgo-Sotelo, B., Torralba, A., & Oliva, A. (2009). Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition*, *17*(6–7), 945–978.

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, *8*(14):18: 1–26, doi:10.1167/8.14.18. [PubMed] [Article]

Follet, B., Le Meur, O., & Baccino, T. (2011). New insights into ambient and focal visual fixations using an automatic classification algorithm. *i-Perception*, *2*(6), 592–610.

Foulsham, T., & Kingstone, A. (2010). Asymmetries in the direction of saccades during perception of scenes and fractals: Effects of image type and image features. *Vision Research*, *50*(8), 779–795.

Foulsham, T., Kingstone, A., & Underwood, G. (2008). Turning the world around: Patterns in saccade direction vary with picture orientation. *Vision Research*, *48*(17), 1777–1790.

Friedrich, T. E., & Elias, L. J. (2014). Behavioural asymmetries on the greyscales task: The influence of native reading direction. *Culture and Brain*, *2*(2), 161–172.

Fuller, J. H. (1996). Eye position and target amplitude effects on human visual saccadic latencies. *Experimental Brain Research*, *109*(3), 457–466.

Garcia-Diaz, A. A., Leborán, V. V., Fdez-Vidal, X. R., & Pardo, X. M. (2012). On the relationship between optical variability, visual saliency, and eye fixations: A computational approach. *Journal of Vision*, *12*(6):17, 1–22, doi:10.1167/12.6.17. [PubMed] [Article]

Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scan paths in visual search. *Visual Cognition*, *14*(4–8), 704–715.

Godwin, H. J., Reichle, E. D., & Menneer, T. (2014). Coarse-to-fine eye movement behavior during visual search. *Psychonomic Bulletin & Review*, *21*(5), 1244–1249.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley.

Greene, M. R., Liu, T., & Wolfe, J. M. (2012). Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns. *Vision Research*, *62*, 1–8.

Harel, J., Koch. C., & Perona, P. (2006). Graph-based visual saliency. InSchölkopf, P. B. Platt, J. C. & Hoffman T. (Eds.), *Advances in neural information processing systems*, (pp. 545–552). Published at the Twentieth Annual Conference on Neural Information Processing Systems, Dec 4–7, 2006, Vancouver, BC, Canada.

Henderson, J. M., Castelhano, M. S., Brockmole, J. R., & Mack, M. (2007). Visual saliency does not account for eye movements during visual 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* (537–562). Oxford, UK: Elsevier.

Jiang, M., Xu, J., & Zhao, Q. (2014). Saliency in crowd. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer vision – ECCV 2014. ECCV 2014. Lecture notes in computer science* (Vol. 8695, pp. 17–32). Cham, Switzerland: Springer.

Judd, T., Durand, F., & Torralba, A. (2012). A benchmark of computational models of saliency to predict human fixations. Technical report MIT-CSAIL-TR-2012-001. Cambridge, MA: MIT Computer Science and Artificial Intelligence Laboratory.

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). Learning to predict where humans look. In *Computer vision, 2009 IEEE 12 international conference on* (pp.2106–2113). Presented at the IEEE 12 International Conference on Computer Vision, Sept 29ndash;Oct 2, 2009, Kyoto, Japan.

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, *10*(4), 346–352.

Koehler, K., Guo, F., Zhang, S., & Eckstein, M. P. (2014). What do saliency models predict? *Journal of Vision*, *14*(3):14, 1–27, doi:10.1167/14.3.14. [PubMed] [Article]

Krüger, H. M., & Hunt, A R. (2013). Inhibition of return across eye and object movements: The role

of prediction. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 735–744.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25), 3559–3565.

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.

Le Meur, O., & Coutrot, A. (2016). Introducing context-dependent and spatially-variant viewing biases in saccadic models. *Vision Research*, *121*, 72–84.

Learmonth, G., Gallagher, A., Gibson, J., Thut, G., & Harvey, M. (2015). Intra-and inter-task reliability of spatial attention measures in pseudoneglect. *PloS one*, *10*(9), e0138379.

Lee, S. P., Badler, J. B., & Badler, N. I. (2002). Eyes alive. *ACM Transactions on Graphics (TOG)—Proceedings of ACM SIGGRAPH 2002*, *21*(3), 637–644.

Loschky, L. C., Larson, A. M., Magliano, J. P., & Smith, T. J. (2015). What would jaws do? The tyranny of film and the relationship between gaze and higher-level narrative film comprehension. *PloS one*, *10*(11), e0142474.

Ma, W. J., Navalpakkam, V., Beck, J. M., Van Den Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nature Neuroscience*, *14*(6), 783–790.

MacInnes, W. J., Hunt, A. R., Hilchey, M. D., & Klein, R. M. (2014). Driving forces in free visual search: An ethology. *Attention, Perception, & Psychophysics*, *76*(2), 280–295.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision*, *11*(8):17, 1–15, doi:10.1167/11.8.17. [PubMed] [Article]

Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLOS Computational Biology*, *8*(2), e1002342.

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, doi:10.1167/8.3.4. [PubMed] [Article]

Nowakowska, A., Clarke, A. D. F., & Hunt, A. R. (2017). Human visual search behaviour is far from ideal. *Proceedings of the Royal Society of London B: Biological Sciences*, *284* (1849), doi: 10.1098/rspb.2016.2767. ISSN 0962-8452. Retrieved from http://rspb.royalsocietypublishing.org/content/284/1849/20162767

Nuthmann, A., & Matthias, E. (2014). Time course of pseudoneglect in scene viewing. *Cortex*, *52*, 113–119.

Ossandón, J. P., Onat, S., & König, P. (2014). Spatial biases in viewing behavior. *Journal of Vision*, *14*(2): 20, 1–26, doi:10.1167/14.2.20. [PubMed] [Article]

Over, E. A. B., Hooge, I. T. C., Vlaskamp, B. N. S., & Erkelens, C. J. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research*, *47*(17), 2272–2280.

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A.-K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2*(2): 4, doi:10.16910/jemr.2.2.4.

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.

Riche, N., Mancas, M., Duvinage, M., Mibulumukini, M., Gosselin, B., & Dutoit, T. (2013). RARE2012: A multi-scale rarity-based saliency detection with its comparative statistical analysis. *Signal Processing: Image Communication*, *28*(6), 642–658.

Smit, A. C., Van Gisbergen, J. A. M., & Cools, A. R. (1987). A parametric analysis of human saccades in different experimental paradigms. *Vision Research*, *27*(10), 1745–1762.

Stainer, M. J., Scott-Brown, K. C., & Tatler, B. W. (2013). Behavioral biases when viewing multiplexed scenes: scene structure and frames of reference for inspection. *Frontiers in Psychology*, *4*, 624.

Takeda, Y., & Yagi, A. (2000). Inhibitory tagging in visual search can be found if search stimuli remain visible. *Perception & Psychophysics*, *62*(5), 927–934.

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, doi: 10.1167/7.14.4. [PubMed] [Article]

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., 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, doi:10.1167/11.5.5. [PubMed] [Article]

Tatler, B. W., & Vincent, B. T. (2008). Systematic

tendencies in scene viewing. *Journal of Eye Movement Research*, *121*, 72–84.

Tatler, B. W., & Vincent, B. T. (2009). The prominence of behavioural biases in eye guidance. *Visual Cognition*, *17*(6–7), 1029–1054.

Torralba, A., Oliva, A., Castelhano, 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.

Tseng, P.-H., Carmi, R., Cameron, I. G. M., Munoz, D. P., & Itti, L. (2009). Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision*, *9*(7):4, 1–16, doi:10.1167/9.7.4. [PubMed] [Article]

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.

Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(2), 145–156.

Viviani, P., Berthoz, A., & Tracey, D. (1997). The curvature of oblique saccades. *Vision Research*, *17*(5), 661–664.

von Wartburg, R., Wurtz, P., Pflugshaupt, T., Nyffeler, T., Luthi, M., & Muri, R. M. (2007). Size matters: Saccades during scene perception. *Perception*, *36*(3), 355–365.

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*(2), 1–16.

Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Yun, K., Peng, Y., Samaras, D., Zelinsky, G. J., & Berg, T. L. (2013). Studying relationships between human gaze, description, and computer vision. In *Computer vision and pattern recognition (CVPR), 2013 IEEE conference on* (pp. 739–746). Presented at the IEEE Conference on Computer Vision and Pattern Recognition, June 23–28, 2013, Portland, OR.

Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, *36*(14), 2177–2187.