Dear Professor Burr,

Thank you for the reviews on our manuscript. We have found both reviewers’ comments helpful and have revised the manuscript to take their suggestions and concerns into account. Replies to the specific points raised by the reviewers are below.

We feel that the manuscript is significantly improved by the changes we have made in response to the reviewers’ comments and we hope that it is now suitable for publication in the *Journal of Vision.*

Yours Sincerely,

Alasdair Clarke

Note: While not asked for by either reviewer, upon returning to this manuscript we decided to slightly change the models used to compare with Flow. Specifically, we have replaced the “re-fit” model with our updated (truncated) version of the central bias presented in Section 2.2. We felt it was an oversight in our original manuscript to introduce this model but then not include it in the flow evaluation section. The comparison with the “re-fit” model was considered to be unnecessary given our previous publication comparing the original central bias model to “re-fit” and we prefer to use this opportunity to evaluate new models.

Reviewer #2 (Comments for the Author (Required)):

“My major concern about this paper is that the proposed "saccadic flow" model has no reference to neurophysiology and it is unclear how this new model improves over the "central bias" model. The authors stated that the saccadic flow model took into account the immediate previous fixation history. However, since the central bias is strong, it is no surprise that saccades are still more likely to land around the center of images regardless of where they are launched from. I believe this is why the saccadic flow model performed only slightly better (or no better) than the central bias model in the gaze landscape analysis and the saliency analysis. To my understanding, the most significant advance brought by the saccadic flow model is that it utilizes fixation history as a prior in predicting the most likely fixation location. In theory this model should be capable of generating more realistic scan-paths as compared to previous models. This aspect of the model, however, was not thoroughly explored in the paper.”

In our revised manuscript we now include measures to clarify the extent to which saccadic flow improves over central bias models. These data are shown in Figures 4 and 5 and discussed in the text. These additional help clarify that Flow does outperform central bias alone. Section 3.3 uses the model to generate scan paths and shows how well they match scan paths generated by human observers.

With regard to the lack of reference to underlying neurophysiology, we acknowledge that a discussion of this was lacking from the original submission and we are grateful for the Reviewer pointing this out. While the model is developed based on empirical data rather than neurophysiology, the fact that it is created from observed saccades means that the statistical descriptions that underlie the model have built in to them the physiological constraints that govern saccade generation. Thus this statistical model encompasses the underlying physiological constraints on saccade generation despite not making these explicit in the development of the model. We now explain this with an additional section in the manuscript as follows:

“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 neurophysiological 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, and Cools [1987], Viviani, Berthoz, and Tracey [1977]. Similarly, the skew in saccade amplitudes toward favouring 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.” [section 2.4.4]

“a) In the Intro section, the authors discussed saccadic momentum and inhibition of return as biases that affect oculomotor behavior. Both of these theoretical concepts predict that saccade landing positions are biased away from multiple previous fixations. The saccadic flow model, however, considered only the immediate preceding fixation. To me, the saccadic flow idea seems like a limited mathematical implementation of these concepts. The authors should at least briefly mention this modeling limitation in the GD section.”

As requested, we have mentioned this in the discussion (section 4.3):

“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 modelling 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 modelling some aspects of human search behaviour [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).”

b) The authors mentioned at various places (including the abstract) that one of the purposes of the present paper was to explore the leftwards bias. The adding of reflected images to the data sets, however, rendered the leftwards bias irrelevant in the analysis. The authors did present relevant modeling results in section 2.3, but it is unclear whether this analysis included reflected images.

We agree that this aspect of our analysis was somewhat lost in the original version of this manuscript and that we did not sufficiently clarify the associated analyses. We have clarified the description in Section 2.3:

“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.” and “The small size of the R^2 in both instances suggests that by ignoring the leftwards bias, we lose little we lose little explanatory power. This brings the advantage of then allowing us to treat everything as symmetrical, which simplifies the model and increases the amount of data available (by mirroring fixations).”

As the leftwards bias accounts for such a small amount of the variance in the unmirrored data (R^2 = 0.01), we conclude that it is not worth taking account of in the saccadic flow model, given this would require a more complicated model and dealing with skewed distributions.

c) A graphical presentation of the saccadic flow idea would be greatly appreciated. To be more specific, it would be helpful to show the fixation likelihood distributions, given a previous fixation at various places in the image. Figure 1 appears to be one such figure, but it is unclear where were the saccades launched from.

Thank you for the suggestion. We now provide three examples of the distributions with the starting point clearly marked in Figure 9.

d) It would be much clearer to present the comparison between the polynomial regression and robust polynomial regression in sections describing the modeling methods.

Agreed, thank you for your suggestion. The comparison between regression methods have been moved.

e) The synthetic scan-path analysis. The authors did not analyze how well the scan-paths generated by the saccadic flow model capture the real scan-paths generated by human observers. The analysis of amplitudes in saccade sequences was relevant. However, the results (Figure 10d) seem to suggest that the saccadic flow model performed worse than the central bias model.

In figure 10c and d, the saccadic flow model (green) is much closer to the human data (blue) than the central bias (red).

f) The results presented in Figure 6 needs elaboration. For instance, what is "the flow: normal model"?

Sorry, this was a hold over from an earlier naming scheme. This should now be fixed.

g) P. 13, section 2.4.3: the author found that "the relative likelihood of scan-paths made during visual search compared to other tasks varies between datasets". This needs elaboration. The modeling results showed that "task has a slight influence over the mean log likelihood for a scan-path". How well do other models (e.g., the central bias model) perform on this analysis? Will the central bias model do a better job in capturing the scan-paths differences across different tasks?

We have added the central bias to this analysis and updated the text accordingly. Both models are similarly insensitive to task. The following has been added to this section:

“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.”

Reviewer #3 (Comments for the Author (Required)):

My main concern is that the entire article remains descriptive. Why don't the authors use statistics to compare prediction between models. The improvement over the Clarke-Tatler model shown in Figure 4 for instance or the probability of fixations occurring in 20% saliency maps should be compared statistically. Maybe bootstrapping might be useful.

We have change the evaluation of the model in several ways. As requested, all evaluations are now carried out with bootstrapping. An additional advantage of using this approach is that we can draw equal sized samples from each dataset, which makes it easier to compare model fits between datasets [as p(data | model) depends directly on the number of points in the data]. We have also moved from showing the raw log likelihood values to showing the difference in log likelihood between using a uniform distribution and the other models. Note, this approach is very similar to giving the AIC values for each model. [AIC takes the number of parameters, k, into account as well as the likelihood. But as n>>k in all of our cases, the effect of k is trivially small.]

We have also taken the suggestion of computing the probability of fixations occurring in the xx% maps onboard, and extended it to carry out a full ROC style analysis for each dataset.

We refrain from carrying out statistical significance testing, as it is not really appropriate for this kind of modelling work (given the large number of datapoints, a trivial difference in model fit will likely give p<0.001). Also see Cumming (2012, Psych. Science). It should be clear from the results in presented in the new version of the manuscript that the flow model improves on the similar central bias model in every dataset we have tested.

The authors demonstrate how their model accounts better for eye movement behavior than existing models. However, the Discussion should focus also on the question what we can learn on the theoretical side. Which parameters might drive saccadic search behavior? The authors address these questions in their Introduction but do not sufficiently clarify how their model provides answers to them.

We have added the following text in section 4.1:

“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. By producing heatmaps that reflect fixation likelihood under the behaviour biases captured in saccadic flow we can produce landscapes that better reflect the image-dependent biases. Using these to base subsequent analysis of scene content at these locations or of differences in fixation behaviour under different tasks allows the researcher to focus analytical efforts on the viewing behaviour 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 behaviour [Tatler and Vincent, 2009]**,** removing these biases appropriately from analyses is important for effective evaluation of changes in behaviour arising from viewed content or behavioural task. Similarly, any attempt to model the involvement of factors such as image salience in eye guidance, should remove image-independent biases from modelling efforts in order to appropriately evaluate the role of any features under investigation. At present removing the central bias from such modelling efforts is widely recognised [Borji and Itti, 2013; Tatler, Hayhoe, Land, and Ballard, 2011], but we have shown in the present work that saccadic flow offers a better explanation of underlying biases in fixation behaviour and therefore using saccadic flow to remove image-independent biases form 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.”

It does not really become clear what the benefit is of the saccadic flow model compared to the central bias model. If the central bias model is a property of saccadic flow why having both model coexisting? In other words, is saccadic flow the winner in the end? Please state this clearly!

Thank you for your suggestion. We have updated our final conclusion as follows:

“While 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 that using saccadic flow were 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, generalisable and should be of use to researchers interested in eye movements in a variety of scene-viewing paradigms.”