A model of selection history in visual attention

Neda Meibodi (meibodi@uni-marburg.de)

Hossein Abbasi (hossein.abbasi@uni-marburg.de)

Anna Schubö (schuboe@uni-marburg.de)

Dominik Endres (dominik.endres@uni-marburg.de)

Department of Psychology, Philipps-University Marburg, Gutenbergstrasse 18,
35032 Marburg, Germany

Abstract

Attention can be biased by the previous learning and experience. We present an algorithmic-level model of this bias in visual attention that predicts quantitatively how bottom-up, top-down and selection history compete to control attention. In the model, the output of saliency maps as bottom-up guidance interacts with a history map that encodes learning effects and a top-down task control to prioritize visual features. We test the model on a reaction-time (RT) data set from the experiment presented in (Feldmann-Wüstefeld, Uengoer, & Schubö, 2015). The model accurately predicts parameters of reaction time distributions from an integrated priority map that is comprised of an optimal, weighted combination of separate maps. Analysis of the weights confirms learning history effects on attention guidance.

Keywords: Visual attention; Selection history; Integrated priority map; Self information maximization; Feature integrated theory; Ex-Gaussian distribution

Introduction

Selective visual attention is a brain function that filters irrelevant sensory inputs to facilitate focusing on relevant items. Bottom-up and top-down mechanisms have traditionally been proposed to control the process of attention guidance. Object saliency and environment features shape the attentional process in a bottom-up manner while the top-down process is mostly controlled by observer intentions and preferences.

In addition to top-down and bottom-up contributions also 'selection history' can play a significant role in guiding attention toward a specific target (Theeuwes, 2019). Selection history (as a third category of attentional deployment) comes into play when an object is emphasized just because of previous attendance in the same context (Awh, Belopolsky, & Theeuwes, 2012). To clarify the distinction between top-down guidance and selection history, Theeuwes argued that selection history is a fast, effortless, and automatic version of attention control while top-down selection is slow, effortful, and controlled (Theeuwes, 2018).

One special form of selection history has been investigated in (Feldmann-Wüstefeld et al., 2015; Kadel, Feldmann-Wüstefeld, & Schubö, 2017; Henare, Kadel, & Schubö, 2020). These studies combined an associative learning task with a visual search task. The result showed that observers attend more to a stimulus which was predictive in the preceding feature discrimination task. Considering to what extent selection history can be suppressed by top-down process, Kadel et al. (2017) tested three different top-down-influenced modes of task preparations such as pretrial task cuing. As their results showed, attentional biases induced by selection history persisted despite the task preparation.

An integrated priority map was proposed by Awh et al. as a theoretical framework to explain how selection history and other factors of attention guidance interact (Awh et al., 2012; Theeuwes, 2019). Priority maps have been successfully employed by many authors (Fecteau & Munoz, 2006; Zelinsky & Bisley, 2015; Klink, Jentgens, & Lorteije, 2014; Todd & Manaligod, 2017; Veale, Hafed, & Yoshida, 2017; Chelazzi et al., 2014) to explain the result of the processes which shape attention. In a review, Klink et al. (2014) summarized how goal-driven and stimulus-driven maps in cortex combine with a value-based map in midbrain. This combination results in a priority map for the frontal eye fields.

Stimulus-driven (bottom-up) models of attention were developed early on (Itti, Koch, & Niebur, 1998). These models tend to ignore the effects of selection history, task or training (Itti & Borji, 2015). Itti et al. (1998) implemented feature integration theory (three feature maps including color, intensity and orientation), winner-take-all, inhibition of return and a normalization method to model visual attention in a bottom-up manner. Veale et al. (2017) validated a neural implementation of Itti's model. In another bottom-up model, Bruce and Tsotsos (2006, 2009) –using self information maximization ($-\log(p(x))$), where x is a feature – proposed a computational model of saliency that is called 'Attention based on Information Maximization (AIM)', because attention is attracted by surprising, i.e. potentially informative, regions of an image. Furthermore, thanks to deep learning advances, there has been recent progress in deep visual saliency models (Borji, 2019).

Beside above mentioned models, Itti and Borji (2015) reviewed more than 50 computational bottom-up models. They also reviewed some computational top-down models. Such models (Navalpakkam & Itti, 2005; Hwang, Higgins, & Pomplun, 2009; Borji, Sihite, & Itti, 2014) are less well researched than saliency models, which might be due to the fact that they require information not available from the stimulus. There are also some models on how bottom-up and top-down work together in attentional guidance (Chikkerur, Serre, Tan, & Poggio, 2010; Kimura et al., 2008). Chikkerur et al. (2010) used a Bayesian framework to explain how a combination of bottom-up and top-down attentional guidance work together in cortex.

Despite substantial progress in building models of attention, there are still many open questions. Selection history has hardly been modeled. One exception is Tseng et al.'s model of the influence of inter-trial priming – a type of selection history effect – on attention guidance (Tseng, Glaser, Caddigan, & Lleras, 2014). They implemented a Ratcliff-type diffusion model (Ratcliff, 1978) for a 2-forced-choice task and showed that the history can affect Ratcliff diffusion model parameters.

In this paper we introduce an algorithmic-level model (in the sense of Marr (1982)) to show how bottom-up, top-down and selection history compete against each other to guide visual attention toward a specific target. By selection history here we mean the effect of learning from previous experience on the current task (see (Feldmann-Wüstefeld et al., 2015; Kadel et al., 2017; Henare et al., 2020)). The model comprises priority maps to integrate goal-driven, saliency-based and history-related biases in a winner-take-all manner. Bottom-up guidance, feature maps and subsequently saliency maps are made based on 'feature integration theory' (Treisman & Gelade, 1980) and 'self information maximization' (AIM) (Bruce & Tsotsos, 2009). To reflect the effect of selection history and learning in the model, a history map contributes to the integrated priority map. Finally, task-relevant information controls the map integration weights that generate predictions for responses and response times. These integration weights are our model for the top-down influences. We test this model on a behavioral database from an experiment by Feldmann-Wüstefeld et al. (2015). The model can predict the reaction time distribution parameters for each participant and also across the experimental groups. To find the best distribution of reaction times, several probability density functions are compared maximizing log-likelihood and the best fitting one – an ex-Gaussian distribution (Matzke & Wagenmakers, 2009)– is used in the model.

Materials and methods

Experiment

The data used in this study comes from the first experiment of Feldmann-Wüstefeld et al. (2015). They investigated the impact of associative learning on covert selective visual attention. The experiment consisted of a 'practice' and a 'main' phase, in which two types of tasks (learning and search) were performed. A central fixation cross was presented on the screen, which was then surrounded by eight different elements on an imaginary circle (Figure 1). 28 participants were divided randomly into 2 different groups, namely 'color group' and 'shape group'. They were first naive about their group membership, but had to learn it on a trial and error basis in the practice phase.

In the 'practice phase', participants had to learn that either color or shape was the response relevant dimension in this learning task (see Figure 1A). Members of the color group had to report the color of the color singleton (blue or green), whereas members of the shape group had to respond to the shape of the shape singleton (triangle or pentagon). They had to use their left hands to press one of two buttons that were placed on the left side of the response pad. Auditory feedback indicated whether they pressed the incorrect key.

In the 'main phase' a second visual search task was added, and participants performed both tasks in random order. In the search task (Figure 1B), all participants had to report the orientation of a line presented inside a diamond shape target. In half of the trials, a response-irrelevant red circle was presented as distractor. Participants used their right hand to press one of two buttons on the right side of the pad to indicate the line orientation (horizontal versus vertical).

The results of this study showed that the history of selection acquired in the learning task affected the participants' performance in the search task. Stimuli that were predictive of the relevant dimension in the learning task biased attention in the visual search task. The authors suggested that the participants' history of either shape or color selection in the practice phase had resulted in a selection history bias.

We presented a model of this selection history bias in the current study based on the behavioral data from the main phase, which comprises at total of 28672 trials across all participants. More details about the experiment can be found in (Feldmann-Wüstefeld et al., 2015).

The Algorithmic Model

Based on the theoretical considerations outlined in the introduction and a preliminary data analysis, we assembled an algorithmic-level model to explain how top-down and bottom-up influences competitively interact with visual selection history to guide attention toward a specific stimulus. The results of this preliminary analysis, that was aimed at determining experimental factors influencing responses and reaction times, are not shown here for space constraints. Inspired by the integrated priority maps in (Awh et al., 2012), we used a 'history map' reflecting the influence of selection history on current attention deployment, see Figure 2. Additionally, there is an overall saliency map for bottom-up influences. How these maps combine into an integrated priority map is controlled by the task in a top-down fashion. Fig-

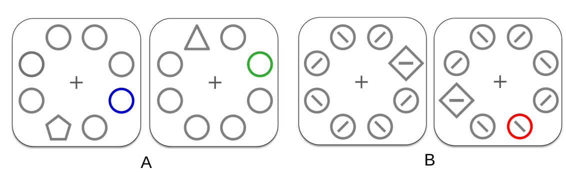

Figure 1: Learning task (A): Participants in the color group had to respond to the color (green vs. blue) and participants in the shape group had to respond to the shape (pentagon vs. triangle). Search task (B): The orientation (horizontal vs. vertical) of the line embedded in the diamond had to be reported. Distractor-absent trial (left). Distractor-present trial (right).

ure 2 also shows how the output of the integrated priority map feeds into a two-part neural network that predicts ex-Gaussian distribution parameters (Luce, 2008) of reaction times (left exit path in the figure) and response likelihoods (the right exit path).

The input stage of the model is based on feature-integration theory (Treisman & Gelade, 1980). The model extracts three types of features (color, shape and orientation) and feature maps –as shown in Figure 2– are computed. In the next processing step, saliency maps that model the effect of bottom-up control on visual attention (Koch & Ullman, 1985) are formed from the feature maps. Shannon's measure of Self-Information is applied, similar to Attention Based on Information Maximization (Bruce & Tsotsos, 2009), to compute saliency maps. Eq (1) and Eq (2) show the actual calculations behind map computation. Feature maps are $M \times N \times K$ vectors where M is the number of trials, N is the number of objects in each trial and K is the number of distinct values that each feature can take on, i.e. we are using 1-out-of-K encoding for the features, with the value 1 indicating which feature value is present. In the current experiment $M = 1024$ (for each participants), $N = 8$ and $K = 4$. Figure 3 illustrates the method of building feature maps for some example trials. For all trials, we take the feature maps f_i for $i \in \{color, shape, orientation\}$ and compute the self-information X_i:

$$\forall k : X_i[k] = -log\left(\sum_{n=1}^{N} f_i[n][k]/N\right) \tag{1}$$

which yields the saliency of all trials $s_i[n]$:

$$\forall n : s_i[n] = X_i\left[\arg\max_k (f_i[n][k])\right] \tag{2}$$

where, due to the 1-of-K feature encoding, we can use *argmax* to pick the self-information corresponding to the current feature value.

Saliency maps s_i are fed into the integrated priority map along with history information (h) to compete in a soft winner-take-all model (Theeuwes, 2019) for the predicted response target. Selection history, the third category of attentional guidance (Awh et al., 2012), carries the effect of learning (participants learned about color or shape in our experiment) into the priority map (p):

$$\forall m,n : p[m][n] = \operatorname{soft}\max_n \left(\sum_i \left(w_{s_i} * s_i[m][n]\right) + w_h * h[m][n] \right) \tag{3}$$

The weights (w_h for history and w_{s_i} for $i \in \{color, shape, orientation\}$) are used to combine the history map and the saliency maps and reflect the effect of the task in a top-down manner. The softmax function is used to ensure that the winning location receives the most attention while keeping the map interpretable as a probability distribution. In our model, Eq 3 can be interpreted as the first layer of a (two-layer) neural network. The second layer is a (linear) mapping from the integrated priority map to reaction time distribution parameters:

$$\forall m : d = \sum_{n=1}^{N} (p[m][n] * w_d) + B_d \tag{4}$$

When w and B are weights and biases of ex-Gaussian distribution parameters' for $d[m] \in (\mu[m], \sigma[m], \tau[m])$.

We also compute a 1-out-of-K representation of the target information ($g[m][n]$ in Eq 5, see also Figure 3) which is used for machine-learning the weights with which the history map and the saliency maps are combined. The weights (w_h, w_{s_i} and w_d) for a task are determined by maximizing the log of the joint distribution of the reaction times (RT), the target g under the distribution predicted by the integrated priority map and the prior distributions over the model parameters δ:

$$\begin{aligned} L &= \sum_{m=1}^{M} \log\left(ExG\left(RT[m] \,|\, \mu[m], \sigma[m], \tau[m]\right)\right) \\ &+ \left(\sum_{m=1}^{M}\sum_{n=1}^{N} \log(p[m][n]) * g[m][n]\right) + \delta \end{aligned} \tag{5}$$

where ExG is ex-Gaussian distribution function. δ is computed as the sum of the logs of the following prior distributions:

$$\begin{aligned} w &\sim \mathcal{N}(0.0, 1.0) \\ B_\mu &\sim \mathcal{N}(600.0, 100.0) \\ B_{\sigma^2} &\sim \mathcal{N}(75.0, 4.0) \\ B_\tau &\sim \mathcal{N}(200.0, 20.0) \end{aligned} \tag{6}$$

Mean and standard deviation of these distributions are selected in a way that matches results from similar experiments (Feldmann-Wüstefeld et al., 2015; Kadel et al., 2017). To find the weights and biases that maximize the joint probability (Eq 5), we draw random initial values from these distributions and then optimize using Python 3.7.6, PyTorch 1.6.0 and Adam optimizer with learning rate 0.2. Code and training data for the models can be found here: `http://dx.doi.org/10.17192/fdr/64.2`

Results and Discussion

To investigate how selection history quantitatively influences attentional guidance, three versions of the model with different history maps are tested. In the first version, the history map contains the response-relevant features in the learning phase (blue and green for the color group, triangle and pentagon for the shape group). In the second version of the model, the history map includes all color singletons (for participants in the color group) and all shape singletons (for participants in the shape group). The assumption is that the participants have learned response-predictiveness on the dimensional level (color or shape), not on the level of single features

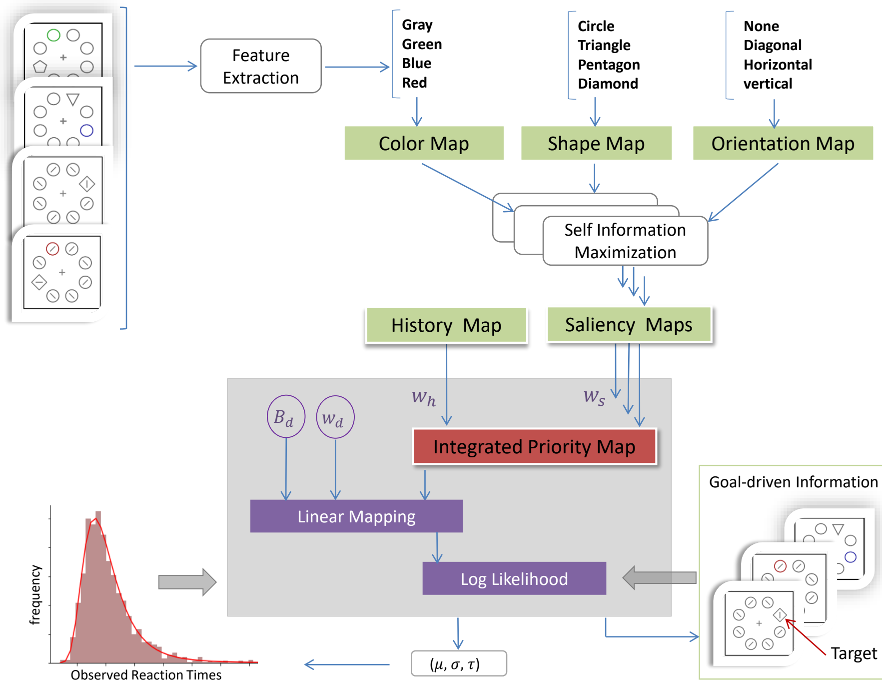

Figure 2: An overview of the algorithmic model. The blue arrows show the direction of data flow from visual input to response and gray arrows show the direction of feedback. w_s, w_h and w_d are map weights. w_s has three elements for color, shape and orientation. w_d has also three elements for distribution parameters(μ, σ, τ). B_d is distribution parameters' bias containing B_μ, B_σ and B_τ.

(such as green or blue). So not only blue, green, triangle and pentagon but also red and diamond are included. In the third version we exclude the history map from the model testing the assumption that only top-down and bottom-up guidance direct attention. To compare these versions of the model, we use a Laplace-approximation. We compute a second-order approximation of the marginal log-probability of the data given the different models' assumptions. We employ these log-probabilites for two Bayesian model comparisons (Bishop, 2006; Barber, 2012; Endres, Chiovetto, & Giese, 2013): fitting one model per participant, and one model per group. In both cases, a model that includes a history map and maps for those features that were predictive during the learning phase is at least 10^{20} as probable as the alternatives. For more details about the model evidences see Figure 4.

Under the assumption that there is a linear mapping from the priority map to the reaction time distribution parameters, the model machine-learns to predict the history map weight (w_h), saliency map weights (w_s) and also the distribution parameters weights and biases (w_d, B_d) (see Figure 2). To compare the weights and also to see how they vary between the color and the shape group see Figure 5, which shows the weights for model version one.

As can be seen in Figure 5, the 'history map' has a higher weight in the color group than in the shape group: to solve the learning task, the color group model has to rely on its learning history features (blue and green) in half of the trials, i.e. in the learning task. Although these colors could be found in the 'color map' as well, there is another color (red) in this map which is task-irrelevant and has to be suppressed. This may be the reason for the increased attention capture by the red distractor in color group members which is reported in (Feldmann-Wüstefeld et al., 2015).

For the search task, a high orientation weight is employed by the color group model, since this task can be solved by spotting an orientation singleton, cf. Figure 1, B.

In contrast, the shape group model can afford to rely less on its 'history map' because the items in its history (triangle and pentagon) exist in the 'shape map' too (triangle, pentagon and diamond), and there is no shape distractor. Therefore, by using a high shape map weight, both the learning task can be solved, and attention can be guided to the shape singleton containing the target in the search task (diamond).

To summarize, the weight of the 'orientation map' is larger in the color group than in the shape group, indicating that the color group model employs orientation saliency in the search task. Using orientation saliency, it does not need to attend to the shape singleton in the search task. However, the shape group model focuses on the 'shape map' which is response-relevant in both tasks.

Also, the weight of the 'color map' was higher in the color group than in the shape group model, since the latter can ig-

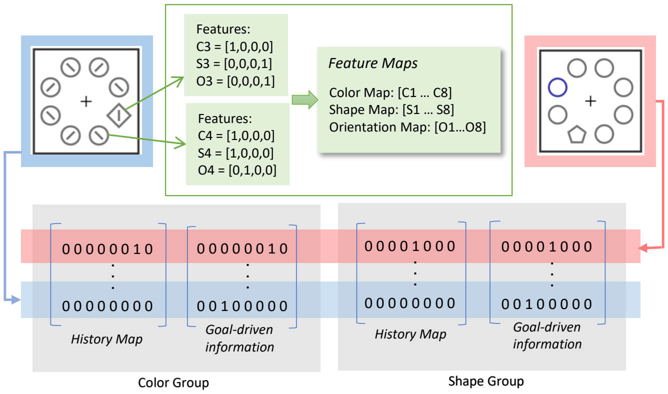

Figure 3: Feature maps, history map and goal-driven information for two random trials. We use 1-out-of-K encoding for the feature vectors, i.e. all components but one are zero. The nonzero component indicates the feature value (see the green box). In each row of history map the location of learned feature is marked. In the target (goal-driven) information the location of response-relevant feature is marked.

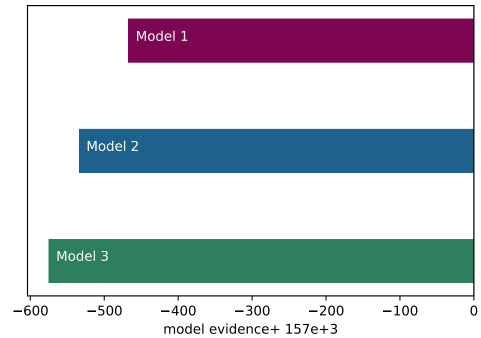

Figure 4: Model comparison. We computed a Laplace-approximation to the Bayesian model evidence across participants. Bigger evidence is better. Model version one, whose history map contains relevant features, scores best. For model descriptions, see text.

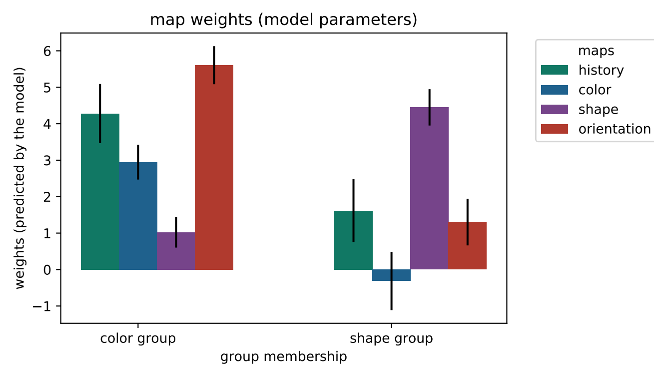

Figure 5: Map weights. For both color group and shape group, optimal map weights for model one are shown. A higher weight means a stronger influence of the corresponding map onto the response and reaction time. The error bars represent the standard deviations of the posterior, i.e. standard errors.

nore color altogether.

The model approximates the reaction time distribution parameters (μ, σ, τ) very well (as can be seen in Figure 6). To quantify how close the model-predicted distributions are to the best fit to the data, we evaluate an approximation to the KullbackLeibler (KL) divergence (Bishop, 2006):

$$KL(p||q) = \int p(RT) \log\left(\frac{p(RT)}{q(RT)}\right) dRT \qquad (7)$$

$$\approx \frac{1}{M}\sum_{m=1}^{M} \log p(RT_m) - \frac{1}{M}\sum_{m=1}^{M} \log q(RT_m)$$

where RT_m is the reaction time in trial m, p(RT) and q(RT) are model-predicted and best-fit distributions respectively. For both color and shape group RTs, we find $KL(p||q) \leq 10^{-4}$ which is very close to the minimal possible value.

Conclusion

We presented a model of selection history in visual attention. The model implements the idea that selection history has a role in attention guidance as claimed by Feldmann-Wüstefeld et al. (2015). We compared different versions of the model and the results show that the one which includes selection history, beside bottom-up and top-down control, is best suited for a quantitative description of the behavioral (RT) results. Our model successfully implements an integrated priority map as proposed by Awh et al. (2012). To determine if this integrated priority map approach is indeed the best description of human behavior, future research needs to investigate non-integrated alternatives. Furthermore, as humans use their attention system in a large variety of situations, a model of task switching needs to be added, rather than training one model per task. The search for such alternatives might be facilitated if we knew what the attentional system is actually trying to achieve on a quantitative level. This is a question situated on the 'computational level' (Marr, 1982). Therefore, we intend to build a computational model in a Bayesian/optimal feedback control framework for both ideal and non-ideal observer-actors. Stochastic evidence accumulation approaches – that have been applied in some other models such as Race Models (Mordkoff & Yantis, 1991) and Drift Diffusion models (Luce, 2008) – might be useful to this end. Another interesting avenue of investigation, which would help in constraining the model, would be the addition of physiological variables. For example, adding EEG signals to disentangle target and related sub-processes (such as enhanced target processing or distractor suppression) would shed further light on attentional guidance processes.

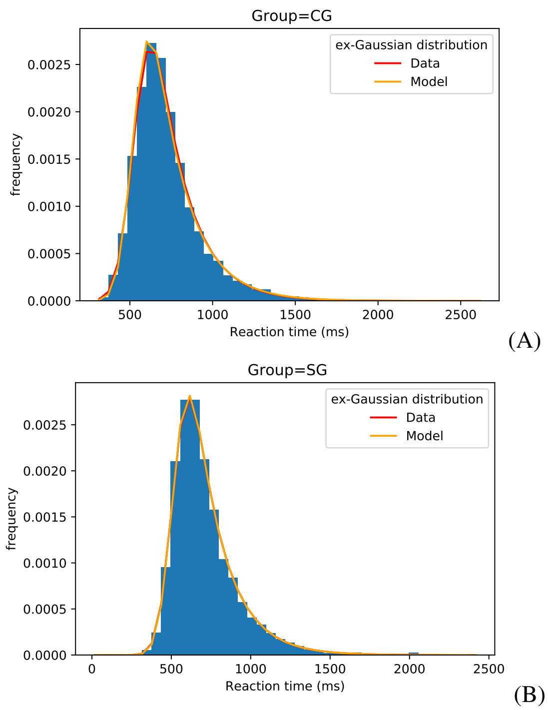

Figure 6: Ex-Gaussian distributions of reaction times. Best fits to the data (red) and model predicted distributions (orange) for participants in (A): color group (CG). (B): shape group (SG).

Acknowledgments

This work was supported by the DFG SFB-TRR 135 "Cardinal Mechanisms of Perception", projects C6 and B3, and "The Adaptive Mind", funded by the Excellence Program of the Hessian Ministry for Science and the Arts. We thank Sara Müller for her contributions to earlier versions of the model.

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. Retrieved from http://dx.doi.org/10.1016/j.tics.2012.06.010

Barber, D. (2012). *Bayesian Reasoning and Machine Learning*. Cambridge: Cambridge University Press. Retrieved from http://ebooks.cambridge.org/ref/id/CBO9780511804779

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Borji, A. (2019). *Saliency Prediction in the Deep Learning Era: Successes, Limitations, and Future Challenges.* Retrieved from https://arxiv.org/abs/1810.03716

Borji, A., Sihite, D. N., & Itti, L. (2014). What/Where to Look Next? Modeling Top-down Visual Attention in Complex Interactive Environments. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *44*(5), 523–538. Retrieved from http://ilab.usc.edu/

Bruce, N. D., & Tsotsos, J. K. (2006). Saliency based on information maximization. *Advances in Neural Information Processing Systems*, *18*, 155–162. Retrieved from http://books.nips.cc/papers/files/nips18/NIPS2005_0081.bib

Bruce, N. D., & Tsotsos, J. K. (2009). Saliency , attention , and visual search: An information theoretic approach. *Journal of Vision*, *9*(3), 1–24. Retrieved from https://doi.org/10.1167/9.3.5

Chelazzi, L., Eštočinová, J., Calletti, R., Gerfo, E. L., Sani, I., Libera, C. D., & Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. *Journal of Neuroscience*, *34*(25), 8594–8604. Retrieved from http://doi/10.1523/JNEUROSCI.0277-14.2014

Chikkerur, S., Serre, T., Tan, C., & Poggio, T. (2010). What and where: A Bayesian inference theory of attention. *Vision Research*, *50*(22), 2233–2247. Retrieved from http://dx.doi.org/10.1016/j.visres.2010.05.013

Endres, D. M., Chiovetto, E., & Giese, M. A. (2013). Model selection for the extraction of movement primitives. *Frontiers in Computational Neuroscience*, *7*, 185. Retrieved from https://doi.org/10.3389/fncom.2013.00185

Fecteau, J. H., & Munoz, D. P. (2006). *Salience, relevance, and firing: a priority map for target selection* (Vol. 10) (No. 8). Retrieved from https://doi.org/10.1016/j.tics.2006.06.011

Feldmann-Wüstefeld, T., Uengoer, M., & Schubö, A. (2015). You see what you have learned. Evidence for an interrelation of associative learning and visual selective attention. *Psychophysiology*, *52*(11), 1483–1497. Retrieved from https://doi.org/10.1111/psyp.12514

Henare, D. T., Kadel, H., & Schubö, A. (2020). Voluntary Control of Task Selection Does Not Eliminate the Impact of Selection History on Attention. *Journal of Cognitive Neuroscience*, *32*(11), 2159–2177. Retrieved from https://doi/abs/10.1162/jocn_a_01609

Hwang, A. D., Higgins, E. C., & Pomplun, M. (2009). A model of top-down attentional control during visual search in complex scenes. *Journal of Vision*, *9*(5), 1–18. Retrieved from https://doi.org/10.1167/9.5.25

Itti, L., & Borji, A. (2015). *Computational models: Bottom-up and top-down aspects.* Retrieved from http://arxiv.org/abs/1510.07748

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliancy-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254 – 1259. Retrieved from https://doi.org/10.1109/34.730558

Kadel, H., Feldmann-Wüstefeld, T., & Schubö, A. (2017). Selection history alters attentional filter settings persistently and beyond top-down control. *Psychophysiology*, *54*(5), 736–754. Retrieved from https://doi.org/10.1111/psyp.12830

Kimura, A., Pang, D., Takeuchi, T., Miyazato, K., Yamato, J., & Kashino, K. (2008). *A stochastic model of human visual attention with a dynamic Bayesian network.* Retrieved from http://arxiv.org/abs/1004.0085

Klink, P. C., Jentgens, P., & Lorteije, J. A. M. (2014). Priority Maps Explain the Roles of Value, Attention, and Salience in Goal-Oriented Behavior. *Journal of Neuroscience*, *34*(42), 13867–13869. Retrieved from https://doi.org/10.1523/JNEUROSCI.3249-14.2014

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–227. Retrieved from https://doi.org/10.1007/978-94-009-3833-5

Luce, R. D. (2008). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York: Oxford University Press.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman and Company.

Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-gaussian and shifted wald parameters: A diffusion model analysis. *Psychonomic Bulletin and Review*, *16*(5), 798–817. Retrieved from https://doi.org/10.3758/PBR.16.5.798%0A

Mordkoff, J. T., & Yantis, S. (1991). An Interactive Race Model of Divided Attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 520–538. Retrieved from https://doi.org/10.1037//0096-1523.17.2.520

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, *45*(2), 205–231. Retrieved from https://doi.org/10.1016/j.visres.2004.07.042

Ratcliff, R. (1978). A Theory of Memory Retreival. *Psychological Review*, *85*(2), 59–108. Retrieved from https://doi/10.1037/0033-295X.85.2.59

Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, *1*(1), 1–15. Retrieved from https://doi.org/10.5334/joc.13

Theeuwes, J. (2019). Goal-driven, stimulus-driven, and history-driven selection. *Current Opinion in Psychology*, *29*, 97–101. Retrieved from https://doi.org/10.1016/j.copsyc.2018.12.024

Todd, R. M., & Manaligod, M. G. (2017). Implicit guidance of attention: The priority state space framework. *Elsevier*, *102*, 121–138. Retrieved from http://dx.doi.org/10.1016/j.cortex.2017.08.001

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. Retrieved from https://doi.org/10.1016/0010-0285(80)90005-5

Tseng, Y. C., Glaser, J. I., Caddigan, E., & Lleras, A. (2014). Modeling the effect of selection history on pop-out visual search. *PLoS ONE*, *9*(3). Retrieved from https://doi.org/10.1371/journal.pone.0089996

Veale, R., Hafed, Z. M., & Yoshida, M. (2017). How is visual salience computed in the brain? Insights from behaviour, neurobiology and modeling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714). Retrieved from https://doi.org/10.1098/rstb.2016.0113

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 154–164. Retrieved from https://doi.org/10.1111/nyas.12606