

Eye movements as a window to cognitive processes

Peter König
U. Osnabrück; U. Medical Center,
Hamburg; Germany
pkoenig@uos.de

Niklas Wilming
U. Osnabrück, Germany; U. Medical
Center, Hamburg, Germany
n.wilming@uke.de

Tim C Kietzmann
U. Osnabrück, Germany
tkietzma@uos.de

Jose P Ossandón
U. Osnabrück; U. Hamburg; Germany
jpossandon@gmail.com

Selim Onat
U. Osnabrück; U. Medical Center,
Hamburg; Germany
sonat@uke.de

Benedikt V Ehinger
U. Osnabrück, Germany
behinger@uos.de

Ricardo R Gameiro
U. Osnabrück, Germany
rramosga@uos.de

Kai Kaspar
U. Osnabrück; U. Cologne; Germany
kkaspar@uni-koeln.de

Eye movement research is a highly active and productive research field. Here we focus on how the embodied nature of eye movements can act as a window to the brain and the mind. In particular, we discuss how conscious perception depends on the trajectory of fixated locations and consequently address how fixation locations are selected. Specifically, we argue that the selection of fixation points during visual exploration can be understood to a large degree based on retinotopically structured models. Yet, these models largely ignore spatiotemporal structure in eye-movement sequences. Explaining spatiotemporal structure in eye-movement trajectories requires an understanding of spatiotemporal properties of the visual sampling process. With this in mind, we discuss the availability of external information to internal inference about causes in the world. We demonstrate that visual foraging is a dynamic process that can be systematically modulated either towards exploration or exploitation. For an analysis at high temporal resolution, we suggest a new method: The renewal density allows the investigation of precise temporal relation of eye movements and other actions like a button press. We conclude with an outlook and propose that eye movement research has reached an appropriate stage and can easily be combined with other research methods to utilize this window to the brain and mind to its fullest.

Keywords: eye movement, eye tracking, saccades, attention, fixation duration, salience, models

Received September 3, 2016 - Published December 1, 2016.

Citation: König, P., Wilming, N., Kietzmann, T.C., Ossandon, J.P., Onat, S., Ehinger, B.V., Gameiro, R.R., & Kaspar, K. (2016). Eye movements as a window to cognitive processes. *Journal of Eye Movement Research*, 9(5):3, 1-16.

DOI: 10.16910/jemr.9.5.3

ISSN: 1995-8692

Introduction

Eye movements are an important part of human behavior and dramatically impact our perceptual life. They are of outstanding quantitative and qualitative importance, and their measurement, commonly called eye tracking, has come a long way. During the times of Huey (1908), Buswell (1935), and Yarbus (1967), measurements of eye movements started with various techniques that faced serious challenges in terms of usability. These ranged from contact lenses with pointers, reflection methods, to search coils within homogeneous magnetic fields. Nowadays, with the advent of optical methods such as eye tracking has become easy and is now used in combination with many other experimental techniques (Carl et al., 2012; Plöchl et al., 2012; Bulea et al., 2013; Reis et al., 2014). A multitude of experiments has addressed properties of eye movements and relevant neuronal circuits (c.f. Trommershäuser et al., 2009; Kowler, 2011; Gegenfurtner, 2016; Rucci & Poletti, 2016). Although many important questions are still unanswered, research of oculomotor function is a mature field with a wealth of results.

In this article, we would like to highlight a specific aspect of oculomotor research: We aim to use eye movements as a window to cognition. The eyes provide extensive information about the presence of objects and agents as well as their spatial relations and dynamic interactions. The ongoing stream of visual information is critical for common tasks like spatial navigation, foraging, avoidance of potential threats, and object manipulation. Therefore, eye movements also reflect internal cognitive processes that are not directly related to external causes but rather internal goals. Thus, eye movements are an integral part of many cognitive processes, and already Yarbus emphasized that eye movements reflect the human thought processes (Yarbus, 1967). Eye movements by default are not an automatic or reflexive action, thus arguably representing the most frequent decision process carried out by the brain (Einhäuser & König, 2010). This has relevant practical consequences. In a standard decision-making experiment, a typical trial would last several seconds in which the experimenter obtains one data point about the decision process under study. In contrast, in a natural unconstrained context, subjects perform saccadic eye movements three to four times a second, which results in a tenfold higher data rate. Thus, in combination with the relevance of visual information for most human

tasks, eye movements should be suitable to monitor most cognitive processes.

Eye movements in the real-world

To understand the decision-making process involved in eye-movement behavior is not only necessary to perform studies in controlled environments, in which the degrees of freedom to move are highly constrained, but it is also necessary to evaluate how eye movements are decided and performed in unconstrained situations in the real world. An example of this is the studies we performed in collaboration with Erich Schneider, the developer of the EyeSeeCam system (Wagner et al., 2006). The EyeSeeCam holds two infrared eye trackers that monitor both eyes. A microcontroller mediates the gaze direction online to a third, pivotable camera that is aligned with the line of sight. A world camera that records the visual field completes the setup. The combination of world camera, gaze direction, and high-resolution gaze camera gives complete information about how and where the eye movements are directed in the world. We used this system to study eye-head coordination in a variety of natural behaviors such as taking a walk, navigating in a train station, or driving a car (Einhäuser et al., 2007, 2009). Watching such movies recorded while someone drives a car on a German motorway on an informal level can tell a rich story (Fig. 1): First, we see that the driver is spotting another moving car on the ramp (Fig. 1A) that is otherwise not particularly salient. Shortly afterwards, at a time the car is entering the motorway and clearly visible on the right, the gaze moves towards the left mirror, presumably checking whether the left lane is free (Fig. 1B). Finally, after the lane switch, the car is very close with salient backlights, but since the black car is no longer relevant, the gaze is already directed straight ahead towards the distance, where some yet-unidentified objects come into view (Fig. 1C). Similar informal observational studies of complex behaviors in the real world provide comparable results, in which eye movements can be explained in terms of simple information-gathering, high-level directives (Land & Hayhoe, 2001; Hayhoe & Ballard, 2005). More quantitative analysis of eye-movements in real-world scenarios reveal which is the precise visual information required in specific tasks, for instance, steering (Land & Lee, 1994), hitting a ball (Land & McLeod, 2000), the bottom-up constrains of free-viewing behavior during walking (Schumann et al., 2008), or how different

visual cues change adaptively their relevance during a task (Jovancevic-Misic & Hayhoe, 2009). Thus, through eye tracking, it is possible to monitor the evolution of complex cognitive processes unfolding within only a few hundreds of milliseconds. Eye movements provide a window to the pacing and the relevant variables of multi-step behavior that otherwise would be very difficult to disentangle.

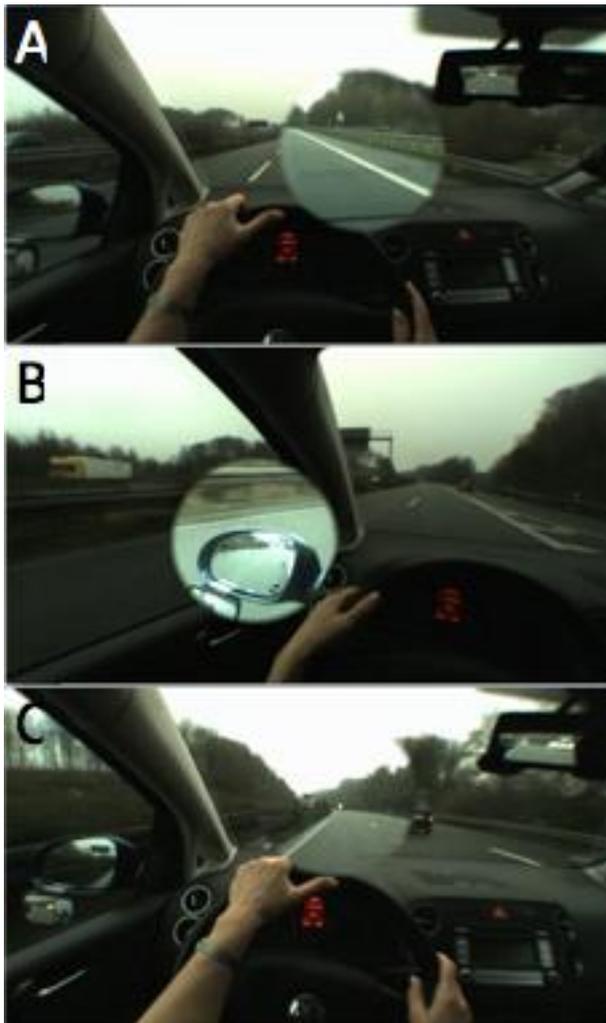


Figure 1. Sequential shots taken by the world camera of the EyeSeeCam system (backdrop) and the gaze camera (circular lens inset with increased contrast) show where attention is directed while driving a car.

Action and perceptual awareness

Can overt attention causally influence our conscious perception? For this, we performed a study in which we

use ambiguous stimuli that may be perceived in different ways. These allow for investigating whether different eye-movement patterns causally influence the resulting percept. Perhaps the best-known example is Necker's cube, described nearly two centuries ago (Necker, 1832). It sparked a debate that soon involved the relation to eye movements (Wheatstone, 1838; Hering 1879). Even today, the question of whether eye movements precede the perceptual switch (Glen, 1940; Kawabata et al., 1978) or are a consequence thereof (Zimmer, 1913; Pheiffer et al., 1956) is unresolved. Here we report on an experiment with 10 sets of ambiguous artistic drawings used as stimuli, including disambiguated versions (Kietzmann et al., 2011; Kietzmann & König, 2015). The example in Figure 2A shows one of these ambiguous stimulus and corresponding disambiguated versions of donkey and seal, where we added or deleted a few line sections. Subjects were naïve to these stimuli, so in contrast to many previous studies including our own, we are not working in a steady state with multiple reversals between recurrent percepts (e.g., Einhäuser et al., 2004). Instead, the subjects view these stimuli for the very first time, and we investigate not multiple reversals, but the first emergence of a percept. Please note that the physical differences between the stimuli are minor. Yet, the resulting eye movement patterns on the disambiguated stimuli are very different (Figure 2B, upper panels). This demonstrates that physically similar stimuli can elicit different patterns of fixation locations. Next, we separate the data of visual exploration of the ambiguous versions according to the reported percept. Comparing those subjects who reported the percept of a donkey and those who reported the percept of a seal reveals a remarkable difference. The former group scanned the ambiguous version much like those subjects who explored the disambiguated donkey stimulus. The latter group, in contrast, explored the ambiguous version much like those subjects who scanned the disambiguated seal stimulus. Remarkably the difference between the two groups exploring the identical ambiguous version is largest at 1300ms before button press. This translates into an effect of about 700ms reaction time corrected, which is a huge effect on a behavioral time scale. This observation holds up for all 10 tested stimulus sets. On average, the eye movements allowed the prediction of the later percept with an accuracy of about 70% (chance 50%). Analyzing the correlation of evidence gathered by subsequent fixations excludes explanations based on evidence-accumulation strategies and additional

experiments demonstrate that manipulating eye movements changes the perceptual outcome (Kietzmann et al., 2011). Thus, this study demonstrates the predictive value of the distribution of fixation locations for the later conscious percept and demonstrates a causal influence of eye movements.

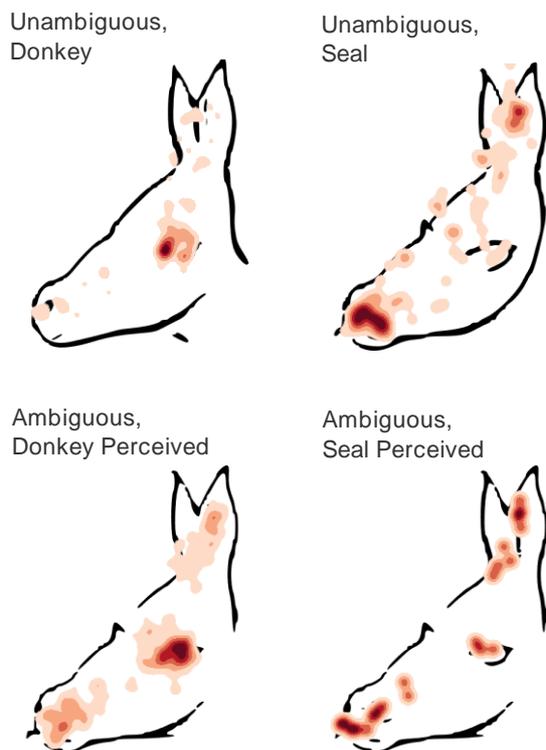


Figure 2. Examples of viewing behavior prior to object awareness on the unambiguous (upper row) and ambiguous (lower row) stimuli with corresponding percepts. There are differences between the groups with different percepts (left and right in lower row), and the differences in the viewing behavior on the ambiguous and unambiguous stimuli are aligned with identical percepts (vertical comparison).

Saliency, maps, and attention

Given the central role of eye movements for action and perception, it is undisputed that many different aspects influence the generation of eye movements: The close relation of eye movements and behavioral goals has been documented in a wide range of contexts (Yarbus, 1967; Land & Lee, 1994; Triesch et al., 2003; Hayhoe & Ballard, 2005; Fecteau & Munoz, 2006; Gozli & Anson, 2016). However, salient events and objects might act as distractors and draw your attention even when they

are unrelated or interfering with the current primary task (Theeuwes, 1991). This stimulus dependent factor can be captured by the concept of a saliency map (Koch & Ullmann, 1985; Itti & Koch, 2001). Several lines of evidence suggest that this bottom-up-directed system is independent from the task-/goal-oriented system (Betz et al., 2010; Pinto et al., 2013) and involves separate neuronal mechanisms (Carrasco, 2011; Corbetta & Shulman, 2002; Desimone & Duncan, 1995; Kastner & Ungerleider, 2000). However, it has also been argued that the dichotomy in top-down and bottom-up control of visual attention is mostly misleading (Awh et al., 2012). A further factor covers spatial aspects like the preference for short saccades (Pelz & Canosa, 2001; Gajewski et al., 2005; Tatler et al., 2006; Gameiro et al., in review), the central bias (Tatler, 2007), the left/right bias (Nuthmann & Matthias, 2014; Ossandon et al., 2014; Kaspar & König, 2011), and saccadic momentum (Smith & Henderson, 2009; Wilming et al., 2013) that all influence the selection of fixation points. This results in a threefold separation of stimulus-dependent (bottom-up), goal-related (top-down), and geometrical factors that jointly control selection of fixation locations (Kollmorgen et al., 2010).

Presently, we focus on the first factor, the stimulus-dependent information. In a bottom-up directed process, different types of image features are analyzed at several spatial scales and integrated into a joint map (Koch & Ullman, 1985). Laurent Itti describes this concept in depth and details the current state of the art (Itti & Koch, 2001; Itti, present volume). Saliency models incorporate a wide variety of visual features like contrast, edges, color, disparity, and motion (Torralba, 2003; Einhäuser & König, 2003; Peters et al., 2005; Baddeley & Tatler, 2008; Frey et al., 2008; Jansen et al., 2009). Therefore, we limit ourselves to some examples and relatively simple models based on regression with visual features (Fig. 3). Testing a variety of visual stimuli (left column), we compare the predicted distribution of fixation locations (central column) with the ground truth data (right column). The hot spots, shown in red, are well captured by our model. However, at an intermediate range the experimental data are sparser than the model's predictions. For quantification, we use the AUC measure. These values capture the correlations of image features and fixation probability that are computed by means of area under the Receiver Operating Characteristic curve. An AUC value quantifies how well fixated and non-fixated image loca-

tions can be discriminated by means of their saliency (Tatler et al., 2005; Wilming et al., 2011). More technically, AUC values quantify the extent to which a certain image feature discriminates between actual and control fixations sampled from other images of the same image category. Thereby, a value of 0.5 indicates random discrimination, and a value of 1.0 indicates perfect predictions. This results in intermediate AUC values. Effectively, the model performs similarly to a prediction of a subject based on the observation of seven other subjects (Wilming et al., 2011). This model is now five years old, and research in this area has advanced quickly. Many parallel developments and later versions of this specific model continue to improve performance (Kümmerer et al., 2015; Huang et al., 2015; Kruthiventi et al., 2015). These models successively reduce the gap to the human-human prediction accuracy baseline, i.e., the upper performance limit of a generic model (Bylinskii et al., 2015). Clearly, saliency-based models have reached a state where they make good predictions with respect to the statistical distribution of fixation locations.

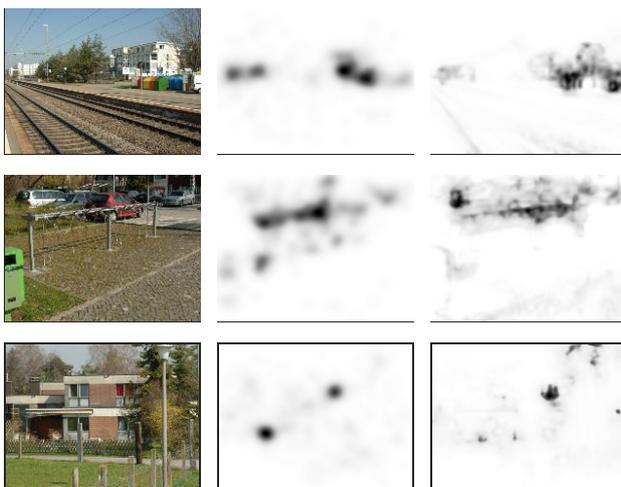


Figure 3. The different columns show a comparison of predictions by a simple saliency model (middle) on a variety of visual stimuli (left) to ground truth experimental data (right).

At this point, we discussed bottom-up factors that influence which fixation locations are selected, and we have to ask: “Are saliency models a mere computational convenience, or is there a real saliency map in the brain?” One way to address this question is the work with neglect patients (Müri et al., 2009). Several months after the lesion, either in the right parietal or frontal lobe, we tested patients in a free-viewing paradigm (Ossandón et al.,

2012). Neglect patients make fewer fixations into the hemifield contralateral to the lesion, here the left visual field. We used a saliency model similar to the one described above to predict fixations in the normal and neglected hemifield of patients and controls. Predictions of fixations in the normal hemifield were as good as the predictions of control subjects’ fixations on either side. However, saliency models based on low-level visual features predicted the fewer fixations in the neglected hemifield even better than those in the healthy hemifield or those performed by the controls. Similar results of increased guidance by low-level features have been reported in other studies with neglect patients (Ptak et al., 2008; Bays et al., 2010; Machner et al., 2012; Fellrath & Ptak, 2015). This observation suggests that the right hemispheric parietal/frontal cortical lesion affected structures mediating top-down directed attention. As a consequence, the presumed low-level saliency map gains increased influence on the selection of fixation locations, i.e., is unmasked. This line of argument makes the assumption that compensatory processes between lesion and time of testing did not affect relevant structures and induced the observations only as a further consequence of the cortical lesion. To test this assumption and further generalize to healthy subjects, we investigated healthy subjects with repetitive transcranial magnetic stimulation (rTMS) (Ossandón et al., 2012). This technique induces a temporary inhibition in the targeted cortical regions, here bilaterally in parietal cortex. After applying rTMS to the healthy subjects, the performance of the model prediction increased. This again is evidence that a (temporary) cortical lesion unmasks a saliency map. Hence, a saliency map in human cortex is not a mere computational convenience, but real.

Assuming the existence of a saliency map, the next obvious question is: where is it? Prime candidates are cortical visual areas (Mazer & Gallant, 2003; Koene & Zhaoping, 2007; Burrows & Moore, 2009; Menon, 2015), parieto-frontal areas described as attentional modules (Corbetta & Shulman, 2002; Bisley & Goldberg, 2010), and subcortical areas like the superior colliculus (Shen & Pare, 2007; Knudsen, 2011) and pulvinar (Robinson & Petersen, 1992). Receptive fields of neurons in these areas match the low-level features used in saliency maps to a surprising degree. Furthermore, these areas are topographically organized, lending themselves rather naturally to the concept of a saliency map (Hubel and Wiesel, 1974). Indeed, based on psychophysical and physiologi-

cal evidence, several studies argue for the existence of a saliency map in primary visual cortex (Zhaoping, 2008; Zhang et al., 2012; Zhaoping 2016). We tested this hypothesis by a combination of fMRI and eye tracking (Betz et al., 2013). We used pink noise for the visual stimuli because it avoids high-level influences like objects, faces, or text. In one of the quadrants of the image, a large patch was increased or reduced in contrast (Fig. 4A). Importantly, compared to baseline, both manipulations increased the saliency and attracted an additional number of fixations (Fig. 4B). Therefore, if V1 represents visual saliency, we would expect that neuronal activity, as characterized indirectly by the fMRI BOLD signal, increases in both cases. The BOLD signal obtained in V1, however, showed a near-perfect linear relation with luminance contrast. The signal was significantly reduced in the case of reduced luminance contrast and significantly increased in the case of increased luminance contrast (Fig. 4C). Furthermore, to investigate more complex representations of a putative saliency map, we applied linear multivariate pattern-classification techniques. However, we could not decode the location of the salient quadrant independent of the type of the contrast modification only at chance level (Betz et al., 2013). Similar statements could be made for V2 and V3. Thus, in these experiments we could not read out salient image locations in these three topographically organized visual areas. These findings suggested that the BOLD activity in early visual cortex (V1–V3) is dominated by contrast-dependent processes and does not include the contrast invariance necessary for the computation of a saliency map.

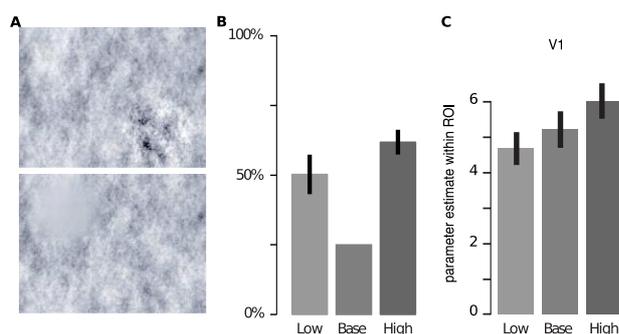


Figure 4. (A) Examples of pink-noise stimuli with increased (top) and decreased (bottom) contrast modifications. (B) Fraction of fixations made to modified quadrant for all three conditions. The saliency of a quadrant is increased by decreased luminance contrast and by increased luminance contrast. (C) Mean BOLD activation in V1 in the three

conditions. A decrease in contrast leads to a decrease in BOLD signal. An increase in contrast leads to an increase in BOLD signal. Error bars represent SEM across subjects.

However, other studies suggest that a saliency map need not be strictly localized, but might be an emergent phenomenon of stepwise processing in the cortical hierarchy (Treue, 2003; Soltani and Koch, 2010). In line with this view, saliency might not necessarily be a consequence of purely visual aspects of an image. Sensory-motor aspects of scenes might also be relevant for fixation selection (Humphreys et al., 2010). For example, tool objects capture more attention than pictures of non-tool objects, highlighting the saliency associated with objects based on how they relate to our bodies. Visual stimuli are generally shown on 2D presentation settings, therefore limiting to a large extent affordances that a visual scene might offer with respect to our bodies. We presented 3D and 2D versions of the same scenes and analyzed binocular visual features at fixated locations (Jansen et al., 2009) using the ground truth depth maps. When pictures were binocularly presented in 3D, the first fixation points were consistently directed towards parts of a scene that were closer to the viewer, suggesting that humans first look at parts of a scene with which they can interact. This highlights the fact that even if stimuli are shown on a simple image plane, body schemes and the relationship of the image shown with respect to our body play an important role.

Sampling and inference

Visual acuity declines dramatically with increasing eccentricity. The information available on upcoming fixation locations drops systematically with increasing saccade amplitude. As a consequence, selecting fixation locations based on saliency shows a tradeoff: On one hand, you may select a close-by location where you have much information available, by a short saccade. Alternatively, you may select a distant location by a long saccade, about which you know little and the average information to be gained is large.

When we plot the target of all saccades while aligning the point of origin, a steep decline in frequency of occurrence is obvious (Fig. 5A, Ossandón et al., unpublished). This decline may be caused in part by the reduced information available on saccadic targets at high eccentricity and in part by properties of the oculomotor system favor-

ing short saccades. However, we can utilize a special property of the blind spot region in the retina. Here the optic nerve leaves the retina, and in a large region of several degrees no photoreceptors are available. As a consequence, under monocular viewing conditions no direct information on the visual scene is available in that region. Instead, a filling-in process interpolates the visual information and leads to a seamless perception. At the blind spot, perception is based on the surrounding information, and therefore no structures can stand out of the surround, i.e., be salient. Thus, perceived saliency at the blind spot region is systematically reduced. The concept of a saliency map predicts under monocular viewing conditions a reduced number of saccades towards the blind spot region. In contrast, for the same reason, the potential amount of information to be gained by a saccade towards the blind spot region is higher than at other locations of equal eccentricity. The concept of maximal information gain predicts under monocular viewing conditions an increased number of saccades towards the blind spot region. However, against the idea of maximal information, previous studies have demonstrated that saccade amplitudes are reduced, instead of increased, during the exploration of high-pass filtered images in which peripheral information can only be accessed through eye movements (Groner et al., 2008; Ossandón et al., 2014). To evaluate the effect of saliency and information gain, we compared the probability of saccades targeting the blind spot region in the temporal visual field with saccades of equal amplitude to the corresponding location in the nasal visual field (Fig. 5B). Taking the difference, we do not observe a systematic bias in either direction. At very high eccentricities, the amount of available data is reduced, and the signal-to-noise ratio drops. Clearly, this analysis turned out differently than expected, and neither the prediction based on a saliency map nor on the maximal information gain is supported by the data.

Let's take a step back and consider how humans handle expectations and surprising information (Horstmann, 2015). Contemporary theories propose that the brain operates constructively and generates probabilistic models, which are continuously tested against reality, i.e., sensory inputs (Clark, 2013). A probabilistic model successfully explains a large range of phenomena like perceptual illusions (Weiss et al., 2002) and the optimal integration of multi-modal signals (Wolpert et al., 1995; Ernst & Banks, 2002; Körding & Wolpert, 2004). In

predictive coding, a popular version of these models, top-down-directed signals implement predictions, and bottom-up-directed processing relays error signals. Crucially, in this framework the estimate of the consequences of agents' actions, like eye-movements, can also serve as predictions. Indeed, there is evidence for predictive coding for passive stimulation (Murray et al., 2002; Summerfield et al., 2008; Alink et al., 2010; Kok et al., 2012). We continued this line of thought and studied whether changes of the visual input produced by eye movements result in predictable signals in the visual system.

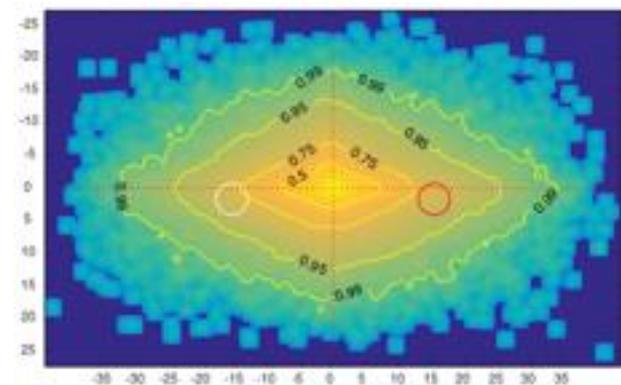


Figure 5. Probability density distribution of fixations during monocular (left-eye patched) free exploration of natural images in a large visual field. Distance from center is in visual degrees, and yellow contours enclose the corresponding probability mass. Red circle indicates the location of human blind spot, and the white circle indicates the corresponding location in the nasal visual field.

Investigating signals that are generated in the absence of actual inputs in the retina's blind spot requires the combination of several techniques (Ehinger et al., 2015). We employed monocular stimulus presentation in the blind spot region of one eye or at the corresponding location of the other eye. With online eye tracking, we implemented a gaze-contingent stimulation. When the subject performed a saccade from the centrally placed fixation spot towards the location of the blind spot (or of equivalent eccentricity) with a specified probability, the display was changed on the fly. For example, a collinear Gabor patch might be changed to a Gabor patch with an annulus aperture and a central orthogonal inset (Fig 6). Thus, based on the previously sampled peripheral vision, there will be a violation of the subject's prediction of what will be visible at the upcoming fixation location. By measuring EEG, we assessed the physiological substrate

related to processing the violation of the subject's expectations.

With this combination of techniques we investigated the relation of inference and predictions across saccades. Aligning the data on stimulus onset, we observed a significant event-related potential of the inside/outside blind spot factor (the weights of the GLM, i.e., discounted for the other factors; Ehinger et al., 2015). This captured the physiological differences of the filling-in process at the blind spot in one eye and bottom-up stimulus processing of stimuli presented at the corresponding location of equal eccentricity at the other eye. Next, we looked at the data aligned to saccadic offset. At this point in time, the gaze focused on the previously eccentric stimulus for in-depth processing. The main factor change/no change captured whether the stimulus has been modified during the saccade or not. We found a significant main effect with a timing and topography compatible to the P3. This

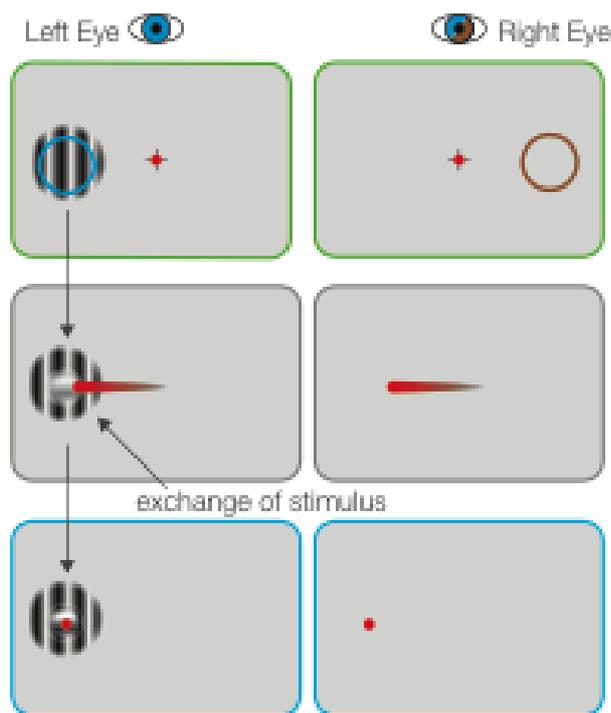


Figure 6. Experimental design for the investigation of violation of expectations based on sampled evidence vs. inference. Each set of two panels shows the stimulus presented to the left and the right eye, respectively. After a fixation interval, a stimulus appeared monocularly in the periphery (top). After the disappearance of the fixated crosshair, the subjects perform a saccade to the center of the presaccadic stimulus (middle). In this example, contingent on the gaze position the central part of

the stimulus is changed by the postsaccadic stimulus (bottom). The colored circles represent the location of the blind spot in each eye and were not displayed on the screen.

demonstrates by electrophysiological means that human subjects perform predictions of what will be in the central visual field after a saccade. The third and final step is crucial. Are violations of predictions based on veridical data or inference processes handled similarly? We observed a significant interaction between inside/outside blind spot factor and the change/no change factor starting about 200ms after saccade offset. In summary, these data demonstrate that the brain treats violations of predictions across saccades differently, depending on whether they are based on directly sampled information or inference.

In combination, the last two studies raise interesting questions. The latter demonstrates that the brain has information on whether information is directly related to sensory input or is based on an inference process. Yet, in the selection of fixation locations, the blind spot region, where information is based on inference, is selected as often as the horizontally mirrored region at equal eccentricity, where the information is based on direct sensory input. The brain has information available on the different origin of the information, but it does not consider that difference in selecting where to sample the visual stimulus next.

Visual context and timing.

Research on eye movements is a highly active research field. In recent years, the number of publications on processes relating to where to fixate next under natural conditions has increased considerable. Yet, the question of when to move on to the next fixation location is neglected in comparison (Nuthmann, et al., 2010; Nuthmann, 2016). In fact, it is related to the idea of ambient and focal processing stages or global to local processing (Unema et al., 2005). Above we argued that each eye movement involves a decision. Thus, analyzing the timing of this process is highly valuable.

At the beginning of this review, we described an experiment on the exploration of ambiguous visual stimuli (Kietzmann et al., 2011). In this study, we also investigated the systematic modulation of the time devoted to local processing relative to the time spent sampling new information, the exploration/exploitation dilemma (Berger-Tal et al., 2014). We took the average fixation dura-

tion as an indicator on the emphasis of exploration vs. exploitation. That is, short fixation durations indicate a bias against local processing and a priority for exploration of the whole stimulus. In contrast, long fixation durations indicate a priority on exploitation of the present fixation. We modulated the available information in a classification task according to a 2x2 design by embedding the ambiguous/disambiguated stimuli in a scene context or presenting them in isolation. Importantly, with the start of the first saccade, the display of context was removed and only the ambiguous stimulus was maintained. Thus, it was impossible to explore the context by saccades, and virtually all fixations were located on the centered ambiguous stimulus. We observed a significant main effect of ambiguity (-27ms, with shorter fixation durations for ambiguous stimuli) and context (+52ms, with longer fixation durations in the context condition), but no significant interaction. Thus, the ambiguity of a stimulus had an effect of moderate size favoring exploration, i.e., sampling information at other locations. The context had a large effect favoring exploitation. These data suggest that, given the additional information supplied by the short perception of the scene context, the classification task is easier, and we cut down on exploration and devoted more time to the in-depth analysis at the location where we fixated. Indeed, the reaction time data support this hypothesis. The ambiguity on average led to a prolongation of reaction time by +439ms. The context, in contrast, induced a reduction of reaction time by -942ms. Please note that the effects on the reaction time

have the opposite sign of the effect on fixation data. With information based on the context, subjects spent more time with local visual analysis but reacted faster. In summary, the time spent with local analysis during a recognition task seems to be precisely controlled and can be modulated by the difficulty and contextual information.

Is there another way to systematically modulate exploration/exploitation of subjects? A direct approach is to vary the amount of information to be explored, i.e., stimulus size (von Wartburg et al., 2007). We presented stimuli full screen on a 30" monitor (Gameiro et al., in preparation). Furthermore, cropped and scaled versions were shown in 7", 10", 15", and 21" size. We analyzed visual exploration in terms of saccadic amplitudes and fixation duration. The mean saccadic amplitude scaled with screen size in a manner that was nearly perfect linearly. A linear scaling of the whole distribution of saccadic length could potentially cause this. However, looking at the distributions of saccadic amplitudes revealed a more complex picture (Fig. 8). The peaks, i.e., the most probable saccade amplitudes, for the 7" and 30" stimuli were very similar. Thus, we followed an alternative approach. We performed a simulation based on the data obtained with 30" stimuli. We sampled saccadic vectors of the 30" condition and discarded probabilistically data according to the spatial bias observed for the 7" stimuli. The resulting distribution of saccadic amplitudes coincided nearly perfectly with the observed data for the 7" condition. Thus, we can explain the distribution of saccadic amplitudes for small stimuli not by a scaling of the distribution, but by a sampling process tied to the spatial bias of the region of interest.

As a next step, we investigated whether we can apply these insights to observed distributions of fixation duration. Compared to the 30" condition, the data for the 7" stimuli were shifted towards longer fixation durations. This is in line with our previous reasoning, that with a reduced size of stimulus less is to be explored and that priority should be given to exploitation. Can we understand these data based on the same sampling principle? Again, based on the 30" condition, we sampled saccadic vectors (and associated fixation durations) and applied a probabilistic selection process constraint by the observed spatial bias of the 7" condition. The resulting simulated distribution did not approximate the 7" well. Instead, they were close to the original 30" condition data. In sum-

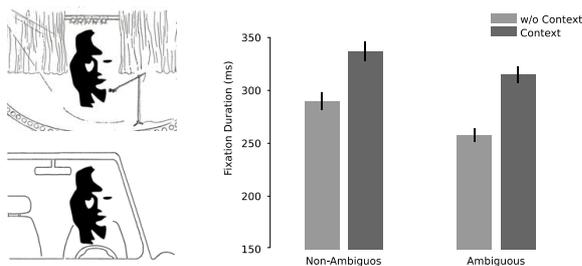


Figure 7. The influence of stimulus ambiguity and context on fixation duration. For every ambiguous stimulus, two contextual scenes were created, which were congruent with either percept of the ambiguous stimulus. In the example it is the silhouette saxophone player during a performance (left top), or a frontal view of a female car driver with light from the right (left bottom). Additionally, disambiguated versions of these stimuli were created. The average fixation duration is dependent on the main factor ambiguity as well as on the main factor context (right). Error bars depict SEM.

mary, the distribution of saccadic amplitudes can be well understood based on a single underlying distribution, dynamically adjusted to the region of interest. However, the distribution of fixation durations reflects the shift from exploration to exploitation and is independently controlled.

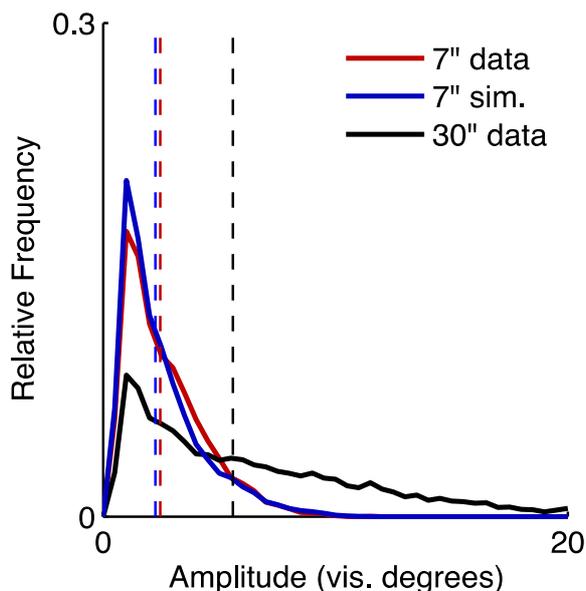


Figure 8. Probability distribution of saccadic amplitudes during free exploration of images of various sizes. The red and black lines show data for the 7" and 30" conditions respectively. The blue lines gives the model prediction for the 7" condition based on the 30" data (for details see text). The dashed lines indicate the median.

Advertising the measurement of eye movements as a window to cognition poses a problem. We make an eye movement about every 200-250ms. For the monitoring of cognitive processes, this temporal resolution is not fully satisfactory. We seek to propose a measure to ease that problem (Kietzmann et al., in preparation). Here, we revisit the experiment in which subjects viewed ambiguous stimuli in a recognition task. Subjects viewed ambiguous stimuli in a recognition task. Specifically, as soon they recognized the stimulus they pressed a button and then verbalized their perception. We observed a broad increase in average fixation duration around the button press peaking just below 400ms (Fig. 9, black line). These surprisingly long fixation durations seem to be in contrast to other data (e.g., Fig. 8). This effect is known as the bus stop paradox (Ito et al., 2003): At every moment in time, we are looking at the average fixation dura-

tion of all ongoing fixations. This introduces a sampling bias towards long fixation durations. Thus, they contribute to more bins than short fixations. This explains why the whole curve is shifted upwards. Alternatively, we considered those fixations that start in a certain interval (Fig. 9, red line). The resulting average fixation duration is considerably lower, matching our expectations. Furthermore, the peak has moved backward in time and peaks more than 200ms before button press. Can we conclude that fixations before the button press are prolonged? This would be jumping to a conclusion, as many of these long fixations last well beyond the button press.

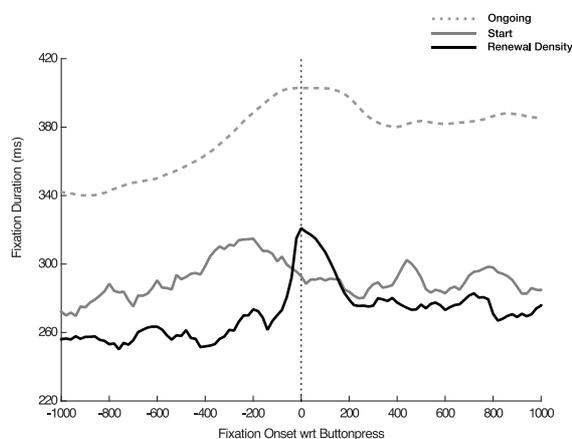


Figure 9. Fixation durations with respect to the button press. The average fixation duration of ongoing saccades is very high, due to the bus stop paradox (grey dashed line). The distribution of fixation duration of starting fixations is shifted left, peaking well before the button press. The simulated fixation duration based on assumed steady state renewal density allows a higher temporal resolution of gaze dynamics (black line).

For a more detailed analysis, we considered the distribution of ongoing fixations at each moment in time. We defined the age of a fixation as the time gone by since the start of the fixation. Next, we computed the total derivative of the density of fixations with respect to time and age. This measure is a hazard function, and for the present purpose we call it renewal density. It gives the fraction of fixations of a specific age that is terminated by a saccade and will give rise to a new fixation at the end of the saccade. Based on this measure, we could calculate the average fixation duration if all fixations would be performed under constant conditions as given at that moment in time. This is in direct analogy to the calculation of average life expectancy. The result showed rather constant fixation duration up until shortly before the

button press (Fig. 9, blue line). Then it rose steeply, peaked at the time of button press, and decayed a bit more slowly afterwards. In this analysis, based on the renewal density the temporal resolution is limited not by the typical fixation duration, but mostly by the amount of data and the ability to align these to a well-defined event.

Conclusion

The investigation of attention and eye movements is a mature field (Groner & Groner, 1989; Pashler & Sutherland, 1998; Pashler, 2016). Yet, the rapid development of eye tracking, the possibility to combine these with experimental techniques, and new methods of data analysis invigorate the interest in research on eye movements.

With the new results on guidance of eye movements and their influence on cognitive processing, it is surprising to find regions of the oculomotor system forming a sparse network within the brain. Instead, we would expect that information related to eye movements is available and processed by many regions, deeply integrated in the cortical connectome, and interacting with cognitive processes. Therefore, in the next 10 years, our view on the guidance and processing of eye movements and the involved brain system might evolve a lot.

Acknowledgements

This research was supported by grants from European Union ERC-2010-AdG-269716 and H2020-FETPROACT-2014, SEP-210141273, 641321 awarded to PK. The authors declare that there is no conflict of interest regarding the publication of this paper.

This review is based on a talk given at the ECEM on August 16, 2015, in Vienna.

References

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, 30(8), 2960–2966.
<http://doi.org/10.1523/JNEUROSCI.3730-10.2010>

- 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):437-43. doi: 10.1016/j.tics.2012.06.010
- Baddeley, R., & Tatler, B. (2006). High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis. *Vision Research*, 46(18), 2824–2833.
<http://doi.org/10.1016/j.visres.2006.02.024>
- Bays, P. M., Singh-Curry, V., Gorgoraptis, N., Driver, J., & Husain, M. (2010). Integration of goal- and stimulus-related visual signals revealed by damage to human parietal cortex. *The Journal of Neuroscience*, 30(17), 5968–5978.
<http://doi.org/10.1523/JNEUROSCI.0997-10.2010>
- Berger-Tal, O., Nathan, J., Meron, E., & Saltz, D. (2014). The exploration-exploitation dilemma: A multidisciplinary framework. *PLoS ONE*, 9(4), e95693–8.
<http://doi.org/10.1371/journal.pone.0095693>
- Betz, T., Kietzmann, T. C., Wilming, N., & König, P. (2010). Investigating task-dependent top-down effects on overt visual attention. *Journal of Vision*, 10(3), 15.1–14. <http://doi.org/10.1167/10.3.15>
- Betz, T., Wilming, N., Bogler, C., Haynes, J.-D., & König, P. (2013). Dissociation between saliency signals and activity in early visual cortex. *Journal of Vision*, 13(14). <http://doi.org/10.1167/13.14.6>
- Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, 33, 1–21.
<http://doi.org/10.1146/annurev-neuro-060909-152823>
- Bulea, T. C., Kilicarslan, A., Ozdemir, R., Paloski, W. H., & Contreras-Vidal, J. L. (2013). Simultaneous scalp electroencephalography (EEG), electromyography (EMG), and whole-body segmental inertial recording for multi-modal neural decoding. *Journal of Visualized Experiments*, (77), 1–13.
<http://doi.org/10.3791/50602>
- Buswell, G. T. (1935). *How people look at pictures*. Location: University of Chicago Press
- Bylinskii, Z., Judd, T., Durand, F., Oliva, A., & Torralba, A. (2015). MIT saliency benchmark.
<http://saliency.mit.edu/>

- Carl, C., König, P., Engel, A. K., & Hipp, J. F. (2012). The saccadic spike artifact in MEG. *NeuroImage*, 59(2), 1657–1667. <http://doi.org/10.1016/j.neuroimage.2011.09.020>
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525. <http://doi.org/10.1016/j.visres.2011.04.012>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03), 181–204. <http://doi.org/10.1017/S0140525X12000477>
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201–215. <http://doi.org/10.1038/nrn755>
- Ehinger, B. V., König, P., & Ossandon, J. P. (2015). Predictions of visual content across eye movements and their modulation by inferred information. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 35(19), 7403–7413. <http://doi.org/10.1523/JNEUROSCI.5114-14.2015>
- Einhäuser, W., & König, P. (2003). Does luminance-contrast contribute to a saliency map for overt visual attention? *The European Journal of Neuroscience*, 17(5), 1089–1097.
- Einhäuser, W., & König, P. (2010). Getting real-sensory processing of natural stimuli. *Current Opinion in Neurobiology*, 20(3), 389–395. <http://doi.org/10.1016/j.conb.2010.03.010>
- Einhäuser, W., Martin, K. A. C., & König, P. (2004). Are switches in perception of the Necker cube related to eye position? *The European Journal of Neuroscience*, 20(10), 2811–2818. <http://doi.org/10.1111/j.1460-9568.2004.03722.x>
- Einhäuser, W., Moeller, G. U., Schumann, F., Conradt, J., Vockeroth, J., Bartl, K., Schneider E, König P (2009). Eye-head coordination during free exploration in human and cat. *Annals of the New York Academy of Sciences*, 1164, 353–366. <http://doi.org/10.1111/j.1749-6632.2008.03709.x>
- Einhäuser, W., Schumann, F., Bardins, S., Bartl, K., Böning, G., Schneider, E., & König, P. (2007). Human eye-head co-ordination in natural exploration. *Network: Computation in Neural Systems*, 18(3), 267–297. <http://doi.org/10.1080/09548980701671094>
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433. <http://doi.org/10.1038/415429a>
- Fellrath, J., & Ptak, R. (2015). The role of visual saliency for the allocation of attention: Evidence from spatial neglect and hemianopia. *Neuropsychologia*, 73(C), 70–81. <http://doi.org/10.1016/j.neuropsychologia.2015.05.003>
- Frey, H. P., Honey, C., & König, P. (2008). What's color got to do with it? The influence of color on visual attention in different categories. *Journal of Vision*, 8(14), 6.1–17. <http://doi.org/10.1167/8.14.6>
- Gajewski, D. A., Pearson, A. M., Mack, M. L., Bartlett, F. N., & Henderson, J. M. (2005). Human gaze control in real world search. *Attention and Performance in Computational Vision*, 3368, 83–99.
- Gegenfurtner, K. R. (2016). The Interaction Between Vision and Eye Movements. *Perception*, 1–25. <http://doi.org/10.1177/0301006616657097>
- Glen, J. S. (1940). Ocular movements in reversibility of perspective. *The Journal of General Psychology*, 23, 243–281.
- Gozli, D. G., & Ansorge, U. (2016). Action selection as a guide for visual attention. *Visual Cognition*, 24(1), 38–50. <http://doi.org/10.1080/13506285.2016.1176095>
- Groner, R., & Groner, M. T. (1989). Attention and Eye-Movement Control - an Overview. *European Archives of Psychiatry and Clinical Neuroscience*, 239(1), 9–16.
- Groner, M. T., Groner, R., & Mühlhelen, von, A. (2008). The effect of spatial frequency content on parameters of eye movements. *Psychological Research*, 72(6), 601–608. <http://doi.org/10.1007/s00426-008-0167-1>
- Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4), 188–194. <http://doi.org/10.1016/j.tics.2005.02.009>
- Hering, E. (1879). Der Raumsinn und die Bewegungen des Auges. *Handbuch der Physiologie*. Vol. III. Part. IV. Leipzig (FCW Vogel).
- Horstmann, G. (2015). The surprise–attention link: a review. *Annals of the New York Academy of Sciences*, 1339(1), 106–115. doi:10.1111/nyas.12679

- Huang, X., Shen, C., Boix, X., & Zhao, Q. (2015). *Sali-con: Reducing the semantic gap in saliency prediction by adapting deep neural networks*. IEEE International Conference on Computer Vision (ICCV), 262 – 270. <http://doi.org/10.1109/ICCV.2015.38>
- Hubel, D. H., & Wiesel, T. N. (1974). Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. *The Journal of Comparative Neurology*, 158(3), 295–305. <http://doi.org/10.1002/cne.901580305>
- Huey, E. B. (1908). *The psychology and pedagogy of reading: With a review of the history of reading and writing and of methods, text, and hygiene of reading*. Location: New York, MacMillan Company.
- Humphreys, G. W., Young Yoon, E., Kumar, S., Lestou, V., Kitadono, K., Roberts, K. L., & Jane Riddoch, M. (2010). Attention and its coupling to action. *British Journal of Psychology* (London, England: 1953), 101(2), 217–219. <http://doi.org/10.1348/000712609X481282>
- Ito, J., Nikolaev, A. R., Luman, M., Aukes, M. F., Nakanatani, C., & Leeuwen, C. V. (2003). Perceptual switching, eye movements, and the bus paradox. *Perception*, 32(6), 681–698. <http://doi.org/10.1068/p5052>
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194–203.
- Jansen, L., Onat, S., & König, P. (2009). Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision*, 9(1), 29.1–19. <http://doi.org/10.1167/9.1.29>
- Jovancevic-Misic, J., & Hayhoe, M. (2009). Adaptive gaze control in natural environments. *Journal of Neuroscience*, 29(19), 6234–6238. <http://doi.org/10.1523/JNEUROSCI.5570-08.2009>
- Kaspar, K., & König, P. (2011). Viewing behavior and the impact of low-level image properties across repeated presentations of complex scenes. *Journal of Vision*, 11(13), 26. <http://doi.org/10.1167/11.13.26>
- Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*, 23, 315–341. <http://doi.org/10.1146/annurev.neuro.23.1.315>
- Kawabata, N., Yamagami, K., & Noakl, M. (1978). Visual fixation points and depth perception. *Vision Research*, 18(7), 853–854.
- Kietzmann, T. C., & König, P. (2015). Effects of contextual information and stimulus ambiguity on overt visual sampling behavior. *Vision Research*, 110, 76–86. <http://doi.org/10.1016/j.visres.2015.02.023>
- Kietzmann, T. C., Geuter, S., & König, P. (2011). Overt visual attention as a causal factor of perceptual awareness. *PLoS ONE*, 6(7), e22614. <http://doi.org/10.1371/journal.pone.0022614>
- Knudsen, E. I. (2011). Control from below: The role of a midbrain network in spatial attention. *The European Journal of Neuroscience*, 33(11), 1961–1972. <http://doi.org/10.1111/j.1460-9568.2011.07696.x>
- Koch, C., & Ullman, S. (1985). Shifts in selective visual-attention - Towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219–227.
- Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, 22(9), 2197–2206. <http://doi.org/10.1093/cercor/bhr310>
- Kollmorgen, S., Nortmann, N., Schröder, S., & König, P. (2010). Influence of low-level stimulus features, task dependent factors, and spatial biases on overt visual attention. *PLoS Computational Biology*, 6(5), e1000791. <http://doi.org/10.1371/journal.pcbi.1000791>
- Kowler, E. (2011). Eye movements: the past 25 years. *Vision Research*, 51(13), 1457–1483. <http://doi.org/10.1016/j.visres.2010.12.014>
- Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244–247. <http://doi.org/10.1038/nature02169>
- Kruthiventi, S. S. S., Ayush, K., & Babu, R. V. (2015, October 10). DeepFix: A fully convolutional neural network for predicting human eye fixations. arXiv.org.
- Kümmerer, M., Theis, L., & Bethge, M. (n.d.). Deep Gaze I: Boosting Saliency Prediction. Pdfs.Semanticscholar.org.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25–26), 1–7. [http://doi.org/10.1016/s0042-6989\(01\)00102-x](http://doi.org/10.1016/s0042-6989(01)00102-x)

- Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, 369(6483), 742–744. <http://doi.org/10.1038/369742a0>
- Land, M. F., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience*, 3(12), 1340–1345. <http://doi.org/10.1038/81887>
- Machner, B., Dorr, M., Sprenger, A., Gablentz, von der, J., Heide, W., Barth, E., & Helmchen, C. (2012). Impact of dynamic bottom-up features and top-down control on the visual exploration of moving real-world scenes in hemispatial neglect. *Neuropsychologia*, 50(10), 2415–2425. <http://doi.org/10.1016/j.neuropsychologia.2012.06.012>
- Mazer, J. A., & Gallant, J. L. (2003). Goal-related activity in V4 during free viewing visual search: Evidence for a ventral stream visual salience map. *Neuron*, 40(6), 1241–1250.
- Menon, V. (2015). Salience network. *Brain mapping*, 2, 597–611. <http://doi.org/10.1016/B978-0-12-397025-1.00052-X>
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences*, 99(23), 15164–15169. <http://doi.org/10.1073/pnas.192579399>
- Müri, R. M., Cazzoli, D., Nyffeler, T., & Pflugshaupt, T. (2009). Visual exploration pattern in hemineglect. *Psychological Research*, 73(2), 147–157. <http://doi.org/10.1007/s00426-008-0204-0>
- Necker, L. A. (1832). LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid. *Philosophical Magazine Series 3*, 1(5), 329–337. <http://doi.org/10.1080/14786443208647909>
- Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review*, 117(2), 382–405. doi:10.1037/a0018924
- Nuthmann, A. (2016). Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review*, 1–24. <http://doi.org/10.3758/s13423-016-1124-4>
- Ossandon, J. P., Onat, S., & König, P. (2014). Spatial biases in viewing behavior. *Journal of Vision*, 14(2), 20–20. <http://doi.org/10.1167/14.2.20>
- Ossandón, J. P., Onat, S., Cazzoli, D., Nyffeler, T., Müri, R., & König, P. (2012). Unmasking the contribution of low-level features to the guidance of attention. *Neuropsychologia*, 50(14), 3478–3487. <http://doi.org/10.1016/j.neuropsychologia.2012.09.043>
- Pashler H. E., Sutherland S. (1998). The psychology of attention, MIT press.
- Pashler H. (2016) Attention, Psychology Press.
- Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, 41(25-26), 3587–3596.
- Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research*, 45(18), 2397–2416. <http://doi.org/10.1016/j.visres.2005.03.019>
- Pheiffer, C. H., Eure, S. B., & Hamilton, C. B. (1956). Reversible figures and eye movements. *The American Journal of Psychology*, 69(3), 452–3. <http://doi.org/10.2307/1419051>
- Pinto, Y., van der Leij, A. R., Sligte, I. G., Lamme, V. A. F., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, 13(3), 16–16. <http://doi.org/10.1167/13.3.16>
- Plöchl, M., Ossandón, J. P., & König, P. (2012). Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, 6, 278. <http://doi.org/10.3389/fnhum.2012.00278>
- Ptak, R., Golay, L., Müri, R. M., & Schneider, A. (2009). Looking left with left neglect: The role of spatial attention when active vision selects local image features for fixation. *Cortex*, 45(10), 1156–1166. <http://doi.org/10.1016/j.cortex.2008.10.001>
- Reis, P. M. R., Hebenstreit, F., Gabsteiger, F., Tschärner, von, V., & Lochmann, M. (2014). Methodological aspects of EEG and body dynamics measurements during motion. *Frontiers in Human Neuroscience*, 8(124), 1–19. <http://doi.org/10.3389/fnhum.2014.00156>

- Robinson, D. L., & Petersen, S. E. (1992). The Pulvinar and visual salience. *Trends in Neurosciences*, 15(4), 127–132.
- Rucci, M., & Poletti, M. (2015). Control and functions of fixational eye movements. *Annual Review of Vision Science*, 1(1), 499–518.
<http://doi.org/10.1146/annurev-vision-082114-035742>
- Schumann, F., Einhäuser-Treyer, W., Vockeroth, J., Bartl, K., Schneider, E., & König, P. (2008). Salient features in gaze-aligned recordings of human visual input during free exploration of natural environments. *Journal of Vision*, 8(14), 12.1–17.
<http://doi.org/10.1167/8.14.12>
- Shen, K., & Pare, M. (2007). Neuronal activity in superior colliculus signals both stimulus identity and saccade goals during visual conjunction search. *Journal of Vision*, 7(5), 15–15. <http://doi.org/10.1167/7.5.15>
- Smith, T. J., & Henderson, J. M. (2011). Does oculomotor inhibition of return influence fixation probability during scene search? *Attention, Perception & Psychophysics*, 73(8), 2384–2398.
<http://doi.org/10.3758/s13414-011-0191-x>
- Soltani, A., & Koch, C. (2010). Visual saliency computations: Mechanisms, constraints, and the effect of feedback. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(38), 12831–12843.
<http://doi.org/10.1523/JNEUROSCI.1517-10.2010>
- Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, 11(9), 1004–1006.
<http://doi.org/10.1038/nn.2163>
- 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.
<http://doi.org/10.1167/7.14.4>
- 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.
<http://doi.org/10.1016/j.visres.2004.09.017>
- Tatler, B. W., Baddeley, R. J., & Vincent, B. T. (2006). The long and the short of it: spatial statistics at fixation vary with saccade amplitude and task. *Vision Research*, 46(12), 1857–1862.
<http://doi.org/10.1016/j.visres.2005.12.005>
- Torrallba, A. (2003). Modeling global scene factors in attention. *Journal of the Optical Society of America, Optics, Image Science, and Vision*, 20(7), 1407–1418.
- Theeuwes, J. (1991). Exogenous and endogenous control of attention - The effect of visual onsets and offsets. *Perception & Psychophysics*, 49(1), 83–90.
- Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology*, 13(4), 428–432.
- Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision*, 3(1), 86–94. <http://doi.org/10.1167/3.1.9>
- Trommershäuser, J., Glimcher, P. W., & Gegenfurtner, K. R. (2009). Visual processing, learning and feedback in the primate eye movement system. *Trends in Neurosciences*, 32(11), 583–590.
<http://doi.org/10.1016/j.tins.2009.07.004>
- 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, 473–494.
- Wagner, Philipp, Bartl, Klaus, Günthner, Wolfgang, Schneider, Erich, Brandt, Thomas, Ulbrich, Heinz (2006). A pivotable head mounted camera system that is aligned by three-dimensional eye movements Proceedings of the Eye Tracking Research and Application Symposium, San Diego, California, USA. pp. 117-124.
<http://doi.acm.org/10.1145/1117309.1117354>
- Wartburg, R. V., Wurtz, P., Pflugshaupt, T., Nyffeler, T., Lüthi, M., & Müri, R. M. (2007). Size matters: Saccades during scene perception. *Perception*, 36(3), 355–365. <http://doi.org/10.1068/p5552>
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6), 598–604. <http://doi.org/10.1038/nn858>
- Wheatstone, C. (1838). Contributions to the physiology of vision. - Part the First. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, 128, 371–394.
<http://doi.org/10.2307/108203?ref=search-gateway:5d7a8329dd0355c8089702c629fc5d72>

- Wilming, N., Betz, T., Kietzmann, T. C., & König, P. (2011). Measures and limits of models of fixation selection. *PLoS ONE*, *6*(9), e24038.
<http://doi.org/10.1371/journal.pone.0024038>
- Wilming, N., Harst, S., Schmidt, N., & König, P. (2013). Saccadic momentum and facilitation of return saccades contribute to an optimal foraging strategy. *PLoS Computational Biology*, *9*(1), e1002871.
<http://doi.org/10.1371/journal.pcbi.1002871.s002>
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882.
- Yarbus, A. L. (1967). Eye movements and vision. Location: New York.
<http://doi.org/10.2307/1410993?ref=search-gateway:98a1ada81caab1e7f99a02797eb4e298>
- Zhang, X., Zhaoping, L., Zhou, T., & Fang, F. (2012). Neural activities in v1 create a bottom-up saliency map. *Neuron*, *73*(1), 183–192.
<http://doi.org/10.1016/j.neuron.2011.10.035>
- Zhaoping, L. (2008). Attention capture by eye of origin singletons even without awareness--A hallmark of a bottom-up saliency map in the primary visual cortex. *Journal of Vision*, *8*(5), 1.1–18.
<http://doi.org/10.1167/8.5.1>
- Zhaoping, L. (2016). From the optic tectum to the primary visual cortex: Migration through evolution of the saliency map for exogenous attentional guidance. *Current Opinion in Neurobiology*, *40*, 94–102.
<http://doi.org/10.1016/j.conb.2016.06.017>
- Zhaoping, L. (2016). From the optic tectum to the primary visual cortex: Migration through evolution of the saliency map for exogenous attentional guidance. *Current Opinion in Neurobiology*, *40*, 94–102.
<http://doi.org/10.1016/j.conb.2016.06.017>