Probing the temporal dynamics of the exploration-exploitation dilemma of eye movements

Benedikt V Ehinger¹, Lilli Kaufhold¹, Peter König^{1,2}

¹ Neurobiopsychology, Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany

² Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

* Correspondence: Benedikt V. Ehinger, Institute of Cognitive Science, University of Osnabrück, Wachsbleiche 27, Osnabrück, 49076, Germany - behinger@uos.de

Abstract

When scanning a visual scene, we are in a constant decision process regarding whether to further exploit the information content at the current fixation or to go on and explore the scene. The balance of these two processes determines the distribution of fixation durations. Using a gaze-contingent paradigm, we experimentally interrupt this process to probe its current state.

Here, we developed a guided viewing task where only a single 3° aperture of an image (“bubble”) is displayed. Subjects had to fixate the bubble for an experimentally controlled time. Then, the previously fixated bubble disappeared, and one to five bubbles emerged at different locations. The subjects freely selected one of these by performing a saccade toward it. By repeating this procedure, the subjects explored the image. We modeled the resulting saccadic reaction times (choice times) from bubble offset to saccade onset using a Bayesian linear mixed model.

We observed an exponential decay between the forced fixation duration and the choice time: Short fixation durations elicited longer choice times. In trials with multiple bubbles, the choice time increased monotonically with the number of possible future targets. Additionally, we found an influence of the saccade amplitude, low-level stimulus properties, and absolute saccade angle on the choice times.

The exponential decay of the choice times suggests that the sampling and processing of the current stimulus were exhausted for long fixation durations, biasing toward faster exploration. This observation also shows that the decision process took into account processing demands at the current fixation location.

Keywords

Fixation durations, exploration-exploitation dilemma, decision processes, gaze-contingent, bubble paradigm

Introduction

Decisions are a central aspect of cognition. Because the world is noisy, evidence of the states of the world needs to be integrated over time (Vickers, 1970). Indeed, physiological studies using the random dot motion paradigm (Gold & Shadlen, 2007; Shadlen, Britten, Newsome, & Movshon, 1996; Shadlen & Newsome, 2001; Smith & Ratcliff, 2004) suggest such an evidence accumulation process. Furthermore, several other decision processes comply with such models. Examples are categorization (Heekeren, Marrett, Bandettini, & Ungerleider, 2004), eye movements (Leach & Carpenter, 2001), or self-initiated button presses (Schurger, Sitt, & Dehaene, 2012). Most commonly, these decision processes are modeled by using a biologically plausible drift-diffusion process (Ratcliff, 2001). The properties and neurobiological mechanisms of decisions are clearly a vast and active research field (Heekeren, Marrett, & Ungerleider, 2008).

While scanning a scene with eye movements, we need to decide when to move our eyes and what target to select for every single saccade. This is arguably the decision with which we are confronted the most often throughout our lives. Eye movements result in significant changes in signals to the brain, and they influence our conscious perception. The obvious way in which they influence our conscious perception is by changing the visual input, but also, more subtly, by the decision about what part of an environment to sample in the future (Kietzmann & König, 2015). A much-investigated example of this decision process can be found in the act of reading: To read this text, you continuously select the target of the next saccade. Extensive data and models have been published in this domain alone (Rayner, 1998, 2009). The decision of where to look next is, of course, not restricted to reading but occurs in all viewing behaviors, e.g., visual search (Najemnik & Geisler, 2005) or free viewing. Eye movements, i.e., the selection of the next fixation points, are evidently prime examples of decision processes.

Two primary decision processes occur in parallel: One decides when to look, the other where to look. A lot of work has been invested in understanding the selection of the next fixation location. Clearly, many different factors contribute to the decision of where to look next (Kollmorgen, Nortmann, Schröder, & König, 2010; König et al., 2016). Most notably, task-dependent factors (Buswell, 1935; Hayhoe & Ballard, 2005; Rothkopf, Ballard, & Hayhoe, 2007), stimulus dependencies (Einhäuser, Spain, & Perona, 2008; Foulsham & Underwood, 2008, 2009; Koehler, Guo, Zhang, & Eckstein, 2014), and geometric dependencies of the trajectory (Henderson & Smith, 2009; Hooke, Over, Van Wezel, & Frens, 2005; Kaspar & König, 2011; Motter & Belky, 1998; Tatler, Baddeley, & Vincent, 2006; Tatler & Vincent, 2009)

exist. For the future fixation location, some of these factors have been summarized in the saliency model (Itti, Koch, & Niebur, 1998; Koch & Ullman, 1985). In recent years, the performance of saliency models has slowly converged to the inter-individual noise ceiling. In other words, models based on features become as good as predictions based on other subjects, and therefore, only inter-individual differences remain to be explained (Bylinskii, Judd, Oliva, Torralba, & Durand, 2016; Kümmerer, Wallis, & Bethge, 2015; Wilming, Betz, Kietzmann, & König, 2011). Furthermore, the concept of a saliency map is not just a computational convenience. Studies investigating neglect patients provide evidence for the existence of a saliency map in humans are more recent additions. Taken together, a quite extensive literature and profound insights exist on the decision of where to look next. (Ossandón et al., 2012), presumably in the superior colliculus (White et al., 2017). Most models estimate saliency based on low-level stimulus properties, such as luminance, contrast, motion, or edges. In addition, high-level object-related features, such as the recall frequencies of objects in a scene, predict eye locations even better (Einhäuser et al., 2008). In recent years, geometric dependencies on the trajectory have increasingly been included in these models. Although spatial bias was used for many years (Tatler & Vincent, 2009), saccadic momentum (Posner & Cohen, 1980; Wilming, Harst, Schmidt, & König, 2013) and horizontal asymmetries (Ossandón, Onat, & König, 2014)

The “when” question, i.e., the decision to initiate a saccade to a new target, determines the time available for the processing of the visual information available at the current fixation. Investigating this ongoing decision process solely based on the distribution of fixation durations is difficult, as selecting a new saccade target and extracting information from the current fixation are temporally overlapping processes (Findlay & Walker, 1999; Henderson & Smith, 2009). Several clever paradigms have been established to dissociate these two factors. In gaze-contingent paradigms, such as scene onset delay or scene quality change paradigms, subjects explore an image, and at a critical fixation, the image is temporarily exchanged with a mask (Henderson & Pierce, 2008; Henderson & Smith, 2009), or visually altered (Henderson, Olejarczyk, Luke, & Schmidt, 2014; Walshe & Nuthmann, 2014). These paradigms show that these changes influence the fixation duration immediately. They support the *direct control* theory (Gould, 1973; Rayner, 1978) of fixation durations. The direct control theory states that the processing difficulty should influence the fixation durations and thus the information content of the current fixation. A lower limit exists for how recent information can still influence the choice of what saccade to perform. The double-step paradigm (Becker & Jürgens, 1979) allows for the establishment of a minimal time of ~80 ms when a saccade can still be reprogrammed (Findlay & Harris, 1984) depending on other factors for example task conditions (Walshe & Nuthmann, 2015). However, not all fixations are under direct control.,

Henderson and Pierce (2008) found a different set of fixation durations where presumably the saccadic program could not be stopped. However, explaining this second set of fixation durations by a pure direct control theory might also be possible (Pannasch, Schulz, & Velichkovsky, 2011). These findings have been computationally modeled by using the CRISP model of fixation durations (Nuthmann, Smith, Engbert, & Henderson, 2010). This two-stage model of saccade initiation contains a labile part, where the saccade can still be stopped, followed by a non-labile stage, where the saccade cannot be interrupted anymore. Nuthmann et al. show that with this model, they can model a wide range of paradigms, including the mentioned scene onset delay paradigms. For unrestricted eye movements, the literature is a bit sparser. The most comprehensive study comes from Nuthmann (2016), where she analyzed unrestricted eye movement data associated with multiple tasks for relevant factors for fixation duration. In another recent study, Einhäuser & Nuthmann analyzed the interactions of the “when” and “where” question in free viewing (Einhäuser & Nuthmann, 2016). In spite of this progress, the available literature for the “when” question is lacking compared with the literature for the “where” question concerning eye movements.

The ongoing foveal processing during a fixation is an example of an *exploitation* process. In contrast, the initiation of saccades and thereby the inspection of the environment exemplifies an *exploration* process (Gameiro, Kaspar, König, Nordholt, & König, 2017). These together establish a complex quite similar to the exploration-exploitation dilemma commonly observed in other disciplines (Berger-Tal, Nathan, Meron, & Saltz, 2014; Cohen, McClure, & Yu, 2007; Daw, O’Doherty, Dayan, Dolan, & Seymour, 2006): It is a dilemma because at a given point in time, we can either exploit or explore, but not both. We are in a continuous decision process between exploring the image and exploiting the current view. It follows that the balance of these two processes determines whether we maintain the current gaze location or initiate a new saccade and consequently the distribution of fixation durations. Compared with the traditional exploration-exploitation dilemma, knowledge gain continues throughout fixation. However, the formulation of eye movements as an exploration/exploitation dilemma places the focus on the ongoing decision process. The distribution of fixation durations is the main outcome of this dilemma.

In this paper, we investigate the exploration-exploitation dilemma in a guided viewing paradigm. During free viewing, experimenters usually do not have direct control over the fixation duration. They can usually bias the distribution but not control it. Of course, causal interpretations require controlled experimental interventions. Therefore, the ideal experimental conditions would allow us to control these factors by discretizing the temporal and spatial aspects of eye viewing behavior. Such a paradigm should allow one to precisely modulate all

parameters of interest, i.e., fixation duration, the number of the possible next fixation targets, the saliency at the current and next locations, and the geometric features of multiple saccades. In this study, we made the first step toward such a paradigm. We used guided viewing, where subjects saw only a small aperture of a scene (a *bubble*) at a time at each fixation (Gosselin & Schyns, 2001; Kollmorgen et al., 2010). Following this, we exchanged the current bubble by up to five different bubbles at other parts of the underlying stimulus. The subjects then selected one of these bubbles via a saccade, and the unselected bubbles were removed. The participants explored the image by repeating this process. At the end of the trial, we used a memory task to probe how well the subjects integrated the small apertures to a form a complete image. The main benefit of this new paradigm is that it allowed us to disentangle the processing phase from the time to select a new fixation location.

We were interested in three predictions that the paradigm allowed us to test. First, it follows from the exploration-exploitation dilemma that saccades should be easier to elicit when the exploitation phase is terminated. Our paradigm allowed us to probe this process: The display time of the fixated bubble was experimentally controlled. With a short display time, the visual system is still in the exploitation phase, and subsequent eye movements should be elicited more slowly. Vice versa, with a long display time, the visual system is ready to, literally, move on, and saccades should be elicited more quickly.

Second, we tested whether saccade planning can be naively thought of as a multiple independent evidence accumulator racing to a fixed threshold. The first to reach the threshold decided the time and place of the saccade. Multiple potential targets, and thus multiple accumulators, should result in a shorter delay to elicit a saccade. In our paradigm, we tested this by introducing an additional experimental manipulation: Instead of using the classical guided viewing paradigm, which operates with a single future target location, we allowed subjects to decide between multiple locations.

Third, we investigated how far these processes depend on the actual information being analyzed. We, therefore, used noise and urban images with low and respectively high information content. A reasonable assumption is that urban images contain more information that needs to be exploited. Thus choice times should generally be longer.

Summarizing, we hypothesized that longer forced fixation times on the current bubble lead to a shorter time to elicit a saccade to the next target. Furthermore, we expected a decrease in the time needed to elicit a saccade with the number of possible future targets. Finally, we

183 hypothesized that the choice time is positively correlated with the processing demand at the
184 current fixation location.
185

Methods

Subjects

In the primary study, 35 subjects participated (18–42 years, mean age: 24 years, eight male, three left handed, 14 left dominant eye). In a second experiment, we recorded 10 additional subjects (18–22 years, mean age: 19.5 years, one male, one left handed, one left dominant eye).

In the first experiment, we excluded five subjects. Four sessions had to stop early on the subjects' request; one other subject did not reliably look at the bubbles. In the second experiment, three subjects were excluded from further analysis. One subject stopped early, and two other subjects had excessive errors in eye tracker calibration and drift corrections. Each subject provided written informed consent. The ethics committee of Osnabrück University approved the experiment.

Apparatus and Recording

We used a 24" LCD monitor (Benq XL2420T) with a screen resolution of 1920x1080 pixels and a refresh rate of 120 Hz for presentation purposes. The participants viewed the screen from a distance of 80 cm. The participant's left eye movements were tracked at 500 Hz by using an EyeLink II system (SR Research Ltd., Mississauga, Ontario, Canada). We used a 13-point calibration with a mean validation error of $<0.5^\circ$ and a maximal validation error of $<1^\circ$. Furthermore, we performed a drift correction before each trial.

Figure 1: A) Experimental Procedure. Subjects foveate a single bubble for the forced fixation time. It is sampled from an exponential distribution with a mean of 300 ms to ensure a flat hazard function. The next step removes the bubble and simultaneously displays one to five new bubbles (in this example three). The dependent variable is the time until a saccade is initiated. After the selection of one bubble, and most often before the saccade ends, the other bubbles are removed. The newly foveated stimulus is shown for a newly sampled forced fixation time. B)

Example stimuli urban and pink noise. C) The sampling points for bubbles following the pseudo-random Poisson disc sampling algorithm. D) A single bubble as seen by the subject.

Procedure

In every trial, the subjects explored a given image using small apertures, inspired by the bubble technique (Gosselin & Schyns, 2001): Each trial started with the display of one random single bubble (a *sub-trial*) visible for an experimentally specified time (forced fixation time) (Figure 1 A). After the forced fixation time, the single bubble display was replaced with one to five new bubbles (one bubble in 51.6%, two in 25.8%, three in 12.9%, four in 6.5%, and five in 3.2% of fixations). This staggering was used to ensure a higher signal-to-noise ratio for the single bubbles. Out of a pool of precomputed bubbles, we randomly selected a set of new bubbles (*choice bubbles*), with a linear bias toward bubbles close to the current one. Thus, this procedure gave preference to bubbles that were closer to the currently displayed bubble. The subjects chose one of the new bubbles by performing a saccade toward it. The time needed to initiate the new saccade was termed the *choice time* and was our primary dependent variable. We used the phrase choice time instead of saccadic reaction time, as several bubbles offered a choice for the subject to make. Thus, the choice time included the time to select the new target as well as the time to initiate a saccade. During this choice time period, the current bubble was switched off; only the future bubbles were visible. After the saccade to the new target, this new target was displayed for the next forced fixation time. Concurrently with the saccade detection, the other, non-chosen bubbles were removed from display. As subjects could in principle saccade to a distant bubble passing over bubbles in between, we used an algorithm based on the last three samples to estimate the saccade velocity. After the detection of a saccade, defined by a velocity of greater than 30°/s, we checked for the closest bubble only after the velocity was lower than 50°/s again. These parameters resulted in the best time-delay/ spatial accuracy tradeoff. In the analyzed data, we observed that the other choice bubbles were removed in the online algorithm after on average 11 ms (90% range of -4 ms to 23 ms) after the end of the saccades as the eye tracker detected offline. Please note that minimizing this time was no easy feat, as the eye tracker had a delay of ~5 ms to deliver the sample. The monitor took an additional ~5 ms to switch to 95% target luminance, and with 120 Hz, we had (on average) ~4 ms (+/-4 ms) for the vertical retrace to start again. We think that an average delay of 11 ms to turn off the non-chosen peripheral stimuli is negligible. In any case, a slower reaction time had an influence only on when the other bubbles were turned off. The selected bubble was constantly displayed until its forced fixation time expired. For the analysis, we used the forced fixation time starting at the end of the saccade as the eye tracker detected offline.

The distribution of forced fixation times followed an exponential function with a mean of 295 ms, which resulted in a constant hazard function. Thus, at any point in time of the trial, the probability that a new bubble would appear was held constant. Consequently, subjects could not anticipate when the currently fixated bubble would disappear and a saccade was necessary.

The subjects saw parts of the image at two times. One was the forced fixation time, when the time subjects foveated a single bubble. The second was the choice time after the forced fixation time when the subjects saw only stimuli in the periphery, and the foveated bubble was removed. The first, the summed-up display time over all fixated bubbles, was on average 6.1 s (95% quantile: [6.0 s 6.2 s]) and was similar to the free viewing conditions used in other experiments. Thus, depending on the forced fixation times in one trial, subjects were presented on average 19.3 bubbles (95% quantile: [11 26], SD of 4.5). The second part, i.e., when stimuli were visible only in the periphery (the summed-up choice times), took on average 3.2 s (95% quantile: [2.7 s 3.9 s]). Taking the first part together with the second, we obtained an average of 9.3 s (95% quantile: [8.9 s 10.0 s]) during which subjects saw the content of the stimuli.

Stimuli

Each subject completed 128 trials in total, but 32 of those trials related to a different research question (they are discussed in “2 AFC Memory-Task”). In half of all trials, we used grayscale urban images (1280*960px). These were photographs of Zürich city and surrounding cities and had been used in several previous studies (Wilming et al., 2011). In the other half, we used grayscale pink noise images with the same luminance distribution as the urban images (SHINE-Toolbox, Willenbockel et al., 2010). The bubbles (Figure 1 B-D) were Gaussian patches with a diameter of 3°, acting as apertures on the underlying image (Kollmorgen et al., 2010). Within each bubble, a fovea filter (Loschky, McConkie, Yang, & Miller, 2005) and a Gaussian filter were applied to imitate the visual acuity of the human eye and smooth the transition to the background. The fovea filter ensured that all information of a bubble could be extracted with a single fixation at the center. That is, there was no information gained by exploring a bubble by repeated saccades to different locations within it. A Poisson disk sampling algorithm placed bubbles pseudo-randomly on each image with a minimum distance constraint (Bridson, 2007). The minimum distance of two bubble centers was set to the bubble radius. Thus, the maximal area overlap of two bubbles was smaller than 15%. In contrast to a random placement from a uniform distribution, this algorithm avoided clusters or holes during sampling (Figure 1C). Successive bubbles were enforced to be without any overlap. We used on average 101 bubble locations (range 91-109) on each image.

Subjects were instructed to look at the center of the bubbles. The distribution of the distance of fixations to the bubble center (Figure 2 A) shows that subjects fixated closely to the center. The dashed green line (Figure 2 B) shows the density of distances when sampling from the Gaussian visibility mask. The distance to zero reflects that if one draws two samples (x and y) from a Gaussian, it is unlikely that both of them will be close to zero. Refixations (red) lead to a distribution that supersedes this density (Figure 2 B). From that, we conclude that subjects targeted the center of the bubbles and performed corrective saccades (see also Kollmorgen et al., 2010 for similar results).

Figure 2 A) 2D density plot of the number of fixations. The center is close to the origin. B) Kernel density of distance to bubble center. Blue/Dotted: first fixation, red/solid: first re-fixation, green/dashed: theoretic distribution based on the Gaussian bubble mask

2-AFC memory Task

After each trial, the participants performed a 2-AFC memory task. We presented two bubbles simultaneously next to each other. One was chosen randomly from the pool of possible bubbles in the previously displayed image but was not necessarily shown during the trial. The other bubble was chosen randomly from any another image. The participants indicated via a button press which of the two bubbles could have been part of the image they saw. There was no time limit. The aim of the task was twofold: We wanted to ensure that subjects processed the image content in some depth. The other reason was to use it as a motivation and distractor task.

Data Analysis

Data Preprocessing

We based our analysis on the premise that subjects fixated on the displayed bubbles ($n = 67454$). Thus, we removed all sub-trials where they did not fixate on the bubble, where the saccade detection algorithm failed, or where the calibration was not good enough to directly detect which bubble was fixated on; We excluded sub-trials, with a fixation outside of the currently displayed bubble during the forced fixation time ($n=11390$, 16% of total) and where there was no direct saccade to the next bubble but an intermediate one in-between bubbles ($n=22613$, 33% of total). Further, we excluded sub-trials, where the planned forced fixation time was different from the observed forced fixation time by an arbitrary threshold of more than 40ms ($n=6147$, 9% of total). Such a discrepancy could arise when the online saccade-detection algorithm detected the saccade earlier or later than the more sophisticated offline saccade detection. We used the observed forced fixation time in all analyses, which is the actual time in which the subjects saw the stimulus. In total, we removed 30185 sub-trials. Thus, all in all, 55.3% sub-trials remained (mean per subject 1062.3, range 806-1283). Sub-trials with more than one fixation inside the currently displayed bubble (25.2% of the remaining bubbles) were kept in the analysis. A reanalysis of choice times of sub-trials with only a single fixation did not change the results.

In order to remove extensive outliers that would strongly influence our analysis, we used an outlier detection algorithm (Leys, Ley, Klein, Bernard, & Licata, 2013) based on three times the median deviation (MAD) distance from the median for each subject (with a constant factor of 1.4826 included; thus, it defaults back to σ in the normal case). This procedure removed an additional 5.6% of the sub-trials and is below the 10-15% recommended criterion (Ratcliff, 1993).

Linear Mixed Model

Fixation durations and, in our case, choice times are dependent on a multitude of factors. These can be categorical (e.g., the type of scene, urban against noise), or in our study most often continuous (e.g., the amplitude of the previous saccade, the angle between saccades, or, in our case, the forced fixation time). Either a repeated measures ANOVA or a linear mixed model could account for the repeated measures of subjects. However, a repeated measures ANOVA does not allow for continuous factors. Thus, we choose a Bayesian linear mixed model with uniform priors on all parameters.

The first set of predictors describes our main experimental manipulations. Two of the predictors are of primary interest in this study: The forced fixation time and the number of bubbles. The interaction between these two factors indicates a possible dependency between

the underlying processes. In addition, we model the categorical stimulus type, urban images, against noise stimuli.

The second set of effects is correlative in nature and relates to the spatial relation of the fixated bubble to the bubble fixated on in the previous and next sub-trial. As we did not experimentally influence saccade trajectory, for example, by enforcing a fixed trajectory of bubbles, these spatial effects are correlative predictors and do not allow conclusions on causal relations. The angle in absolute monitor coordinates is a circular predictor. In order to model this dependency, we used a Fourier decomposition with one and two periods for the 360° of the predictors (thus, four predictors in total, including two sines and two cosines). We modeled this effect once from the previous to the current bubble and once from the current to the next bubble. Also, we used the difference of the angles between the previous and the current bubble and the angle between the current and the next bubble. We also included the distances of the bubbles. To account for a central spatial bias, we included the standardized z-transformed x- and y-position and the squared, z-transformed x- and y-position of the bubbles in order to account for a quadratic decay toward the edges of the image, which is usually observed with spatial biases. Also, we model a term to describe the distance of a bubble to the center of the screen.

The third set of predictors incorporated temporal dependencies and sequential influences between the trials and bubble presentations. We used trial number, previous forced fixation time, and previous choice time as predictors.

Bayesian Model Fit

We analyzed the data using hierarchical logistic mixed effects models fitted by the No-U-Turn-Sampler in STAN (B. Carpenter et al., 2017; Homan & Gelman, 2014). For the model specification, we followed the implementation by Sorensen (Sorensen, Hohenstein, & Vasisht, 2016). In maximum likelihood linear modeling terms, all within-subject effects were modeled using random slopes clustered in subject and a random-intercept for subject. We estimate all covariances between the random effects with an LKJ-Prior ($\nu=2$), which slightly emphasizes the diagonal over the off-diagonal. We used treatment coding for all categorical factors and interpret the coefficients accordingly.

We used six MCMC-chains with 1000 iterations each, with 50% of the iterations used for the warmup period. We visually confirmed convergence through autocorrelation functions and trace plots. Furthermore, we calculated the scale reduction factors (Gelman et al., 2013) and ensured the recommended criterion for convergence ($R_{\text{hat}} < 1.1$). To control for an adequate

model fit, we calculated 1000 posterior predictive draws and plotted the median and 95th percentile together with the raw data. The posterior predictions matched the data well, and the model seems to be adequate for our inferences. When displaying the data and posterior predictions in their quantile ranges, the posterior predictive checks showed that our model did not capture all features of the data but merely the central tendency. We find a mismatch in the 2.5% quantile and 97.5% quantile. There, the raw data have higher choice times than the posterior predictive. This results from the skewed distribution of choice times on the subject level (Figure 3). But, most importantly, the posterior predictive median and mean value fit the data appropriately.

Reported Statistics

For Bayesian posterior predictive checks, we report 95% credible intervals (CI) using the Cousineau correction for grouped data (Cousineau, 2005). For posterior parameter estimates, we report 95% posterior CI. For other reported data, we use 95% bias-corrected and accelerated bootstrapped confidence intervals of the mean with Cousineau correction for grouped data where applicable (Cousineau, 2005). Cousineau correction adjusts confidence intervals of repeated or grouped estimates by first subtracting the total average of each subject before calculating the variance.

Software

Experimental software was written using PyGame (<http://pygame.org>). We processed data in MatLab (The MathWorks, Inc., Natick, Massachusetts, United States), Python (<http://python.org>), and R (R Core Team, 2013). We acknowledge the use of other packages used to analyze the data (Bates, 2010; Carr, Lewin-Koh, & Maechler, 2015; Pinheiro, Bates, DebRoy, Sarkar, & R-Core-Team, 2016; van der Walt, Colbert, & Varoquaux, 2011; Wickham, 2007, 2015). All experimental scripts, analysis scripts, and data are publically available under <http://osf.io/ba2pn>.

Results

In this study, we used a new paradigm to break down the free viewing paradigm into controlled subprocesses. While several experimental paradigms exist to experimentally manipulate fixation durations, here, we directly control them using forced fixation periods. Furthermore, we controlled the participants' saccade trajectories with the use of small bubble-like stimuli that act as an aperture on the underlying image. These bubble stimuli allowed us to guide the participants in their exploration.

Figure 3 Kernel density estimates of choice times for each subject. The horizontal line depicts the mean of the subjectwise-median choice times.

Thirty-five subjects viewed 96 images each (plus 32 images intermixed with other trials, see Results 2AFC-Task for a description), and we analyzed 35105 fixations based on the guided viewing paradigm. The mean of the median choice time, that is, the duration until subjects initiated the next saccade, was 156.7 ms (95 percentile range: [123.1 ms, 210.3 ms], Figure 3). As expected for fixation durations and reaction times, our choice time distributions were skewed toward the right tail. We used a Bayesian linear mixed model to explain the variations of the choice time. The units of the linear model parameters were chosen to be intuitive and to facilitate comparison between the effect sizes. We either used the average values of the predictors (246 ms based on the average fixation duration in parts of a big free viewing data set (Wilming et al., 2017)) or maximal values (the total number of trials, 96 from this experiment plus 32 intermixed ones, resulting in 128 trials and 19 bubbles per trial average). We additionally grouped the predictors into three distinct classes: Experimental factors, correlative factors, and sequential effects. Experimental factors are controlled, randomized, and balanced. Correlative factors were analyzed based on the subjects' behavior. The sequential

effects capture possible influences due to previous trials. We present the posterior marginal density results in Figure 4.

Figure 4 Linear mixed model parameter estimates and Bayesian 95% posterior credibility intervals. We state the parameters in intuitive units (third column) to aid interpretation. A 'max' depicts the highest observed value of this factors, and 'avg' the average of independent data (Wilming et al., 2017).

Our first prediction relates to the exploration-exploitation dilemma: Sampling the decision process early after fixation onset should result in longer reaction times than eliciting a new saccade after prolonged initial fixation. All subjects showed an exponential decline in the choice time (Figure 5 AB). For short forced fixation times below 100 ms, we observed choice times of around 175 ms. In contrast, in the case of long forced fixation times, choice times saturate around 150 ms. This compressive nonlinearity has an instantaneous rate of growth of -34.0 [-38.3, -29.6] (95% credibility interval, see Methods) times the forced fixation time normalized by an average fixation duration in free viewing images of 246 ms. We find an

additional linear effect of forced fixation time with a slope of 0.7 [-0.1, 1.5]. This effect is small (a change of 0.7 ms in 246 ms, or a ~3ms change in 1000ms) and the parameter estimate contains 0. Therefore, we do not discuss the linear part of the forced fixation time effect further. Descriptively, we observe an increase in the choice times for forced fixation times larger than 1s (not shown). Due to the exponentially distributed forced fixation times, on average, only 2.4% of the forced fixation times are larger than 1s. Therefore, we do not feel confident in making estimates for values larger than 1s, which are unlikely to influence the model estimates. From the mixed model results, it is clear that the forced fixation time increases the reaction time: the longer the current fixation lasted the faster subjects responded.

Figure 5 Main effect of forced fixation time. Subjects viewed the bubbles for an experimentally controlled time, depicted on the abscissa. We record the time until they choose to perform a saccade to the next bubble on the ordinate. A) blue: data with bootstrapped 95% confidence interval of the mean using Cousineau correction. Red: posterior predictive distribution with 95% credibility interval. B) Single subject kernel density estimates of the data.

The second experimental manipulation is the number of bubbles. After fixating a stimulus for the forced fixation time, the stimulus disappeared and one to five bubbles emerged in the periphery. The subjects then decided on one and performed a saccade onto it. The other bubbles then disappeared before fixation onset. Here, we observed a monotonic logarithmical increase of the choice time from one to five bubbles (Figure 6), with an effect size of 25.1 ms [21.0, 29.2] over five bubbles. While the logarithmical effect captured the data well, a categorical or piecewise-linear component (e.g. one bubble is unique, 2 to 5 bubbles are linear) seemed to capture the data equally well. Theoretic considerations for a possibly higher number of bubbles in future experiments make the logarithm the most reasonable choice with the best generalization. It is evident from the data that there is a monotonic increase in the choice time: the more bubbles that are available for a decision, the longer it takes for one to choose a bubble.

Figure 6: We varied the number of bubbles from which subjects could choose. A) Main effect number of bubbles, blue: data with bootstrapped 95% confidence interval of the mean using Cousineau correction, red: posterior predictive distribution with 95% credibility interval. B) Distributions for individual subjects. Each subject showed an increase over the number of bubbles.

Next, we analyze the interaction between the forced fixation time and the number of bubbles (Figure 7) on the choice time. Here we find an interaction between the number of bubbles and the forced fixation time of -14.7 $[-19.0, -10.5]$. In Figure 7, we smoothed the display to get a more intuitive understanding of this relationship. An example to understand the effect size for this log-log interaction is helpful. The predicted difference between a forced fixation time of 100 ms and 1000 ms for a single bubble is 14.7 ms; for five bubbles, it is 20.0 ms. Thus, in this case, the interaction boosts the single bubble forced fixation effect by 40%. We acknowledge the small absolute effect size. Nevertheless, the credibility interval for this effect is far removed from 0. The interaction shows that the effects of the forced fixation and the number of bubbles on the choicetime are dependent on each other: The more bubbles are available to choose from, the larger the effect of the current forced fixation time.

Figure 7: A) The main effects of the number of bubbles (columns) and the forced fixation time (abscissa). The upper row depicts the posterior predictive with 95% Cousineau credibility intervals, and the lower row depicts raw data with 95% bootstrapped mean with Cousineau correction. B) We show the interaction between the forced fixation time and the number of bubbles. The data from A, lower row, were further smoothed for this plot (generalized additive model with thin plate regression splines). And the mean of each curve was removed. This eliminates the main effect of the number of bubbles. A sharper decline results in trials with many bubbles to choose from compared to trials with a low number of bubbles.

The last experimental effect is the image category. The bubbles of one trial were drawn from urban images in half of the trials and pink-noise images in the other half of the trials. We observed a main effect of 4.3 ms [3.1 ms, 5.5 ms], which means urban bubbles have a longer choice time than noise bubbles. This main effect is surprisingly small, given the vast differences in image statistics between the two categories.

We observed various correlative effects (Figure 8). These effects were not experimentally controlled or balanced and thus are correlative in nature. The first effect is the angle between the currently fixated bubble and the next bubble (Figure 8 A, inset). We modeled this effect using four parameters describing the sine and cosine of the base and two times the base frequency. This procedure allowed us to capture the dynamics of the circular nature of this effect (Figure 8 A). The shape of the marginal effect display follows a smooth sinusoidal curve: Reaction times of saccades toward the upper part of the screen seem to be faster than toward the lower part of the display. But if we look at the model estimate, the maximal average difference (the amplitude), and thus the effect sizes of this summed curve, is only 0.9 ms [0.4 ms, 1.5 ms]. The mismatch between the marginal display and the parameter effect size is due to other predictors explaining a large share of the variance of the effect. We observed a similarly small effect of the angle between the previously fixated bubble and the currently fixated bubble on the decision process at the currently fixated bubble. Here the maximal

524 difference over angles is 0.7 ms [0.3 ms, 1.1 ms]. This effect follows a double-u shape, with
525 increased choice times for vertical saccades and faster choice times for horizontal saccades.

526
527 The difference between the two angles is shown in Figure 8 B and is commonly referred to as
528 saccadic momentum. In line with other research (Wilming et al., 2013), return saccades
529 appear slower than forward saccades, but in our study, this is visible only in the marginal plot.
530 Taking into account other factors, the effect disappears. Whereas in free viewing we observed
531 effect sizes of around 45 ms, here we did not see a reliable effect, with on average 0.4 ms [-
532 0.9 ms, 1.7 ms] for the maximal difference of 0° to 180° (see Discussion). The logarithmic
533 distance to the next bubble (Figure 8 C) predicted the choice time with a maximal effect size
534 of 2.6 ms [0.0 ms, 5.3 ms] between a theoretical saccade of 0.2° and the maximum of 30°.
535 Similarly, it is unlikely that there exists an effect of the distance to the last bubble, with an
536 estimate of -1.3 ms [-3.3 ms, 0.9 ms] for the maximal range.

537
538 Next, we discuss the predictors for the absolute position of fixation. Without these predictors,
539 the aforementioned nonsignificant effects all were significant (although small, model results
540 not shown here). We modeled the absolute position of the chosen (thus future) bubble using
541 a linear and a quadratic term to capture the symmetric and quadratic nature of this spatial
542 bias. The horizontal position had a linear influence of -3 ms [-4.6 ms, -1.5 ms] for the maximal
543 range. The vertical position had a bigger effect of -8 ms [-10.0 ms, -5.9 ms] for the maximal
544 range (which in absolute measures is smaller than the horizontal position). We also observed
545 quadratic effects for the horizontal position of -7.8 ms [-12.8 ms, -2.9 ms] and for the vertical
546 position of -9.0 ms [-12.5 ms, -5.7 ms]. In addition to the absolute position, we also observed
547 a linear distance-to-center effect of -5.1 ms [-8.0 ms, -1.8 ms]. These effects indicate a quicker
548 reaction time the farther away from the center one's eyes rest.

Figure 8: Geometric predictors. A) Angle to the next bubble. B) Difference in angle between two bubbles, over termed saccadic momentum. C) Distance to the next bubble. D, E) Horizontal and vertical position of the bubble that has been selected (not the current foveated one), the distance of the selected bubble to the center of the image. C-F) Degrees depict visual angle.

The last effects we modeled are sequential effects: trial number, bubble number, previous choice time, and previous forced fixation time. The trial number influenced the choice time by 2.9 ms [0.9 ms, 4.8 ms] throughout the 128 trials of the experiment (Figure 9 A). The bubble number describes the n-th bubble a subject saw in each trial. It had, on average, an influence of 2.5 ms [1.5 ms, 3.5 ms] (Figure 9 B) throughout one image, i.e., on average over 19 bubbles. The previous choice time had an influence of -1.9 ms [-3.0 ms, -0.8 ms] for an average fixation time of 246 ms. The previous forced fixation time did not seem to influence (0.01 ms [-0.2 ms, 0.3 ms]) the current choice time. All these effects are considered small.

Figure 9: A) Trial number. In total, 128 trials were used, but 32 intermixed trials were only related to the memory task. B) On average, 19 bubbles were shown in a single trial; this effect depicts a very small increase of reaction time during one trial.

This concludes the correlative effects. Compared to the experimental effects, the effect sizes are very small.

2-AFC Memory Task

Subjects performed a 2-AFC task after completing the sequence of bubbles based on one image. They viewed one bubble taken from the previously shown image (but not necessarily a bubble shown during the previous trial) and another from a different image. If subjects process and integrate bubbles to a whole, they should do well in this task. Performance at chance level would indicate a lack of integration and that each bubble was judged independently from all others.

For two noise stimuli, performance was 52.7% (CI_{95} 49.7%–56.0%, SD : 9.4%), as to be expected at chance level. For trials with one noise bubble and one urban bubble, performance was close to perfect, with 97.4% (CI_{95} 95.8%–98.2, SD : 3.3%). For trials wherein two urban bubbles were compared in the memory task, subjects showed an average performance of 72.8% (CI_{95} 69.0%–76.0%, SD : 10.6%).

It is not clear at what level we can argue that integration occurs. We, therefore, analyzed data of a static image condition, where integration to a whole image can be expected. In this condition, a bubbled version of the image was visible, which consisted of 21 nonoverlapping bubbles. The number was slightly higher than the on average 19 bubbles in the current study. The image was visible for 6 s, a similar time to the current study. 16 noise and 16 urban images were shown intermixed with the sequential trials. Here we see very similar results: For two noise stimuli, performance was 53.6% (CI_{95} 47.9%–60.6%, SD : 18.4%). For trials with one noise bubble and one urban bubble, performance was at 96.5% (CI_{95} 93.8%–98.1%, SD : 6.2%). For trials with two urban images, performance was at 75.6% (CI_{95} 70.7%–80.6%, SD : 15.1%). As the confidence intervals are highly overlapping, it is reasonable to conclude that performances do not differ and subjects did indeed integrate the sequential bubbles to a whole image.

Second Experiment: Internal Replication and Context

In light of the recent replication crisis in the (cognitive) psychological sciences subjects (Aarts et al., 2015; Pashler & Wagenmakers, 2012), we strived to internally replicate our findings on an independent set of subjects. In addition, we were interested in what way context shown at the beginning of a trial influenced the choice times and integration performance. We performed a second experiment ($n = 10$) where we briefly flashed the entire scene for 92 ms (11 displayed frames at 120 Hz) at the beginning of half of the trials. This allowed subjects to extract the gist of the scene (gist condition) but did not allow making any saccades. It has been shown that at around 100 ms, subjects already extract the main features of a scene (Potter, 1976), even when the image was masked, which we did not do here. In the other half of the trials, the

experiment remained identical to the original experiment to allow for within-subject comparisons. The two types of trials were randomly intermixed.

As can be seen in figure 10, we replicated all effects we found in the first experiment. In addition, we did not observe an interaction with the gist of a scene in any of the factors. In the 2-AFC memory task, subjects performed at very similar levels: The case of two noise stimuli was not different from chance level (50.3%, CI_{boot95} : 44.5%–59.3%). One noise and one urban bubble resulted in near perfect performance (98.3%, CI_{boot95} : 97.1%–99.2%). And we found the same results as in the previous experiment. When subjects had to choose between two urban bubbles, there was no improvement in either gist condition (no-gist: 78.8%, CI_{boot95} : 70.4%–84.2%, and with-gist 70.5%, CI_{boot95} : 55%–79.0%). Nevertheless, the second experiment replicates all of our choice time results in a new cohort of subjects. Furthermore, briefly presenting the gist of a scene did not influence either the choice times or performance in the memory task.

Figure 10: The upmost interval of each posterior predictive triplet depicts parameter estimates when subjects did not see the gist of a scene (bright color). The middle interval (dark color) shows parameter estimates when the gist was shown for 92 ms before exploration started ($n = 10$, paired). The lowest interval (gray) depicts the parameter estimate shown in Figure 4 (different subjects, $n = 35$), wherein the gist was not seen either. Showing the gist does not induce any relevant difference in the within-subjects comparison or in comparison with the cohort of the previous experiment.

Discussion

Summary

The distribution of fixation durations can be described by an exploration-exploitation decision process. Here, we used a guided viewing paradigm to control the decision processes occurring during a fixation and dissociated it from the processing of the fixated location. We found an exponential decrease of the time needed to choose the next fixation target dependent on the time available for processing the stimulus at the current fixation location. This dependence provides evidence for the exploration-exploitation dilemma in the decision process. Secondly, we found a monotonic increase in choice time with the number of available saccade locations. These data indicate that potential future saccade targets are accumulating evidence in their favor in a dependent manner.

Exploration-Exploitation

Our first main result directly tested the exploration-exploitation idea: We described eye movements as an ongoing decision process between further exploitation of the current view and further exploration of new, unseen elements. In this study, we interrupted the subjects during the exploitation stage of the current bubble at unpredictable points in time and investigated the influence on the choice time. From the literature, we expected saccadic planning times to be between 100 and 175 ms (Rayner, 1998; Schall & Thompson, 1999), which is in line with the present observation of ~150 ms. Please note that the choice time was measured from stimulus offset. It was not measured from the previous saccade offset, as at that time, no targets for further saccades were available. Thus, the intersaccade time amounts to an average of ~450 ms. In principle, the choice time (starting with stimulus offset and target onset) could be constant and independent of the forced fixation time, ending with stimulus offset and target onset. Instead, we observed an exponential relationship of the choice time and the forced fixation time. This demonstrates that the choice time is dependent on the degree of processing at the current fixation location and gives support for the exploration-exploitation view.

Are There Alternative Explanations for the Exponential Decay?

In simple reaction-time experiments, Drazin (1961) observed an exponential decay when the foreperiod, the period between warning signal and go signal, was increased. This mimics our forced fixation effect and could be an alternative explanation. As discussed in Niemi and Näätänen (1981), the important factor influencing this effect is the predictability of the stimulus. The hazard function, the instantaneous probability that a foreperiod/forced fixation is ending,

is essential here. With a nonuniform hazard function, as in Drazin, subjects can predict with higher certainty when a stimulus is going to appear, whereas when a nonpredictable, flat hazard rate is used (Baumeister & Joubert, 1969; Mowbray, 1964), no effect of foreperiod on the resulting reaction time can be found. We used a uniform, thus unpredictable, hazard function. Therefore, our effect cannot be explained by predictability effects of reaction time.

Related Paradigms

A paradigm that has a very similar procedure to the current study is the masked onset paradigm (Glaholt, Rayner, & Reingold, 2012; Rayner, Smith, Malcolm, & Henderson, 2009). Their goal was to measure the minimal time viewers need to see a scene at each fixation in order to obtain normal viewing behavior. In order to test that, Glaholt et al. masked the stimulus after a given time (50 ms, 75 ms, and 100 ms). This stops any further incoming information. Either the whole stimulus or a circular part around the current fixation was masked. They measured the time until subjects started the next fixation. They found that while in reading, only 50-60 ms of visible stimulus is needed (Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981); in free viewing, information about the current fixation needs to be seen for at least 100 ms. Our results indicate that even more time is needed (at least ~400 ms) until the effect of forced fixation time saturates for all fixations. Furthermore, Glaholt et al. (2012) found a bimodal distribution of fixation durations, which they explain by a saccadic inhibition mechanisms due to the rapid onset of the stimulus mask. Saccadic inhibition (e.g. Reingold & Stampe, 2002) is a delay of saccade production that the onset of a (possibly irrelevant) stimulus has. Its most salient feature is a bimodal distribution in the fixation durations; shortly after the inhibitory stimulus fewer saccades are generated. In our case we did not observe this bimodality in the choice times (Figure 3) or when adding the forced fixation times to the choice times (not shown). Thus, saccadic inhibition cannot explain our results.

Another phenomenon related to the one reported here could be seen in the stimulus onset delay paradigm (Vaughan, 1982; Vaughan & Graefe, 1977). In this paradigm, subjects looked back and forth between two fixation points in the search for a target. At each fixation there was a variable stimulus onset delay, replacing the fixation point either with a target or with a distractor. The time to elicit the next saccade was the dependent variable and can be seen mirrored in our choice time. Similarly, the stimulus onset delay is mirrored in the forced fixation time. If the stimulus was shown with a delay of 300 ms after onset, Vaughan found decreasing response times by 100-150 ms. Thus, the delay speeds up the response time to the appearing stimulus. Here, we observed a substantial yet in comparison small effect of ~30 ms in the first 300 ms. Due to the large difference, it is unlikely that the same mechanisms underlie both effects. The crucial difference between the two tasks is that in our paradigm the stimulus is

directly visible at fixation onset, then subsequently vanishes and the new target appears. Whereas in the stimulus onset task, the stimulus is not visible upon fixation onset but becomes visible after the forced non-stimulation period. We conclude that it seems unlikely that the effect observed due to the stimulus onset delay is related to the exploration-exploitation effect we investigate in the present study.

Evidence Accumulation With Multiple Targets

Our second prediction relates to the evidence accumulation between multiple targets. We assumed one independent evidence accumulator for each target that race to a fixed threshold. The first to reach the threshold is selected. Our observed monotonic increase with the number of future target locations is incompatible with this explanation. This result is compatible with data and evidence accumulator models from Leach and Carpenter (2001). To reconcile these data with such an evidence accumulator model, interactions between the integrators of different locations are necessary. This might occur through different thresholds for multiple future locations as can be found in monkey LIP neurons that decreased drift rates through lateral inhibition due to limitations in sampling capacities (R. H. S. Carpenter & Williams, 1995; Churchland, Kiani, & Shadlen, 2008; Leach & Carpenter, 2001) or more complex interactions with time-varying thresholds (Ludwig, 2009). A future analysis step with this paradigm is to fit drift-diffusion models taking into account the effects described here.

We observed a logarithmic increase between the number of possible targets and reaction times, which is similar to the “Hick’s”-effect (Hick, 1952; Proctor & Schneider, 2017). But there is evidence that saccades do not follow this rule: the “anti-Hicks” effect (Lawrence, 2010; Lawrence, St John, Abrams, & Snyder, 2008). This anti-Hicks effect is commonly observed when there are multiple possible target locations to choose from. These authors differentiate between exogenous (e.g., prosaccades) and endogenous (e.g., antisaccades), where an anti-Hicks effect can be found in the former but not the later. We cannot replicate this finding here. It is well possible that selecting between multiple possible targets is an intrinsic task and our results are in agreement with earlier findings. We think there are two additional important differences: One is that subjects in our study could not predict how many bubbles would appear compared to a blocked design in Lawrence. A second difference is that in our study there is no correct stimulus, whereas in Lawrence there was always one correct target. We could reconcile this by a new experiment with a modified version of our paradigm presenting multiple bubbles but highlighting one as a target. In this case, an anti-Hicks effect could be observed.

Dividing fixation durations in two processes, one starting a new saccade and one staying at the current fixation, has been made explicit in visual search (Beintema, van Loon, & van den Berg, 2005; Van Den Berg & Van Loon, 2005). In their model, both processes are explicitly modeled as two dependent integrators racing to individual thresholds that decide whether to continue exploiting the current view or go on and explore the scene. In visual search experiments, it seems as though the processing of the peripheral stimulus is secondary to the processing difficulty at fixation (Hooge & Erkelens, 1999; Wu & Kowler, 2013). The conclusion from these studies is that in visual search subjects do not bias fixation durations for better target selection, but only for discrimination of the target at hand. This is not to say that the peripheral information is ignored; it is still used for target selection, it just shows a weak influence on the fixation durations. It seems that the subsequent integration of foveal and peripheral information can occur independently (Ludwig, Davies, & Eckstein, 2014), but not always so (Van Den Berg & Van Loon, 2005; VanRullen, Reddy, & Koch, 2004). Contrary to our initial prediction, we found only a very small effect of foveal processing in our experiment: a 4 ms difference between urban and pink noise images. On the other hand, the peripheral decision task showed a difference of up to 25 ms depending on the number of bubbles. Thus, in our experiment, it seems we have a reverse relation to observations in visual search (Hooge & Erkelens, 1999). This clearly shows how strong task dependency can influence the integration of peripheral and foveal processing.

From our finding of a monotonic increase followed a new prediction for free-viewing paradigms. In free viewing, some parts of an image are preferably gazed upon. This can be quantified by empirical saliency, usually measured by the density maps of fixations over multiple observers. It has been recently shown that locations with higher saliency have higher fixation durations (Einhäuser & Nuthmann, 2016). Given our findings and what has been found about the interaction of saliency and fixation durations, we predicted prolonged fixation durations if an image has multiple distinct peaks of empirical salience against other locations without these distinct peaks. The peaks in saliency mimic multiple bubbles as shown in our study.

Computational Models of Fixation Duration

Models of fixation durations were proposed in the subject literature (Engbert, Nuthmann, Richter, & Kliegl, 2005; Reichle, Pollatsek, Fisher, & Rayner, 1998). In recent years, fixation duration models for free viewing have emerged as well; for example, the CRISP (Nuthmann et al., 2010) and ICAT (Trukenbrod & Engbert, 2014) models. In these two models, fixation durations are modeled by two main components, a stochastic random walk for the timing of saccades and one for the saccade generation. For the saccade generation, either a random

draw from a gamma distribution (CRISP) or another stochastic random walk (ICAT) was used. Our data could be used as a benchmark to test the adequacy of these models, especially the saccade generation part of the model. Here we expected a strong modulation of the initial starting point (or the threshold, conversely) of the random walk or of the drift speed of saccade generation. In a first step, a simplified model using drift-diffusion modeling could be appropriate. It is important that with such models not only the average choice times can be modeled, but the whole distribution. Differences between experimental manipulations could be caused, or be hidden in the shape of the response distributions (for a Bayesian hierarchical solution based on Weibull-functions see Rouder et al., 2005) and not captured by the central tendency measurement. The discrete notion of exploitation (fixation on a stimulus) and exploration (which stimulus to select) can help differentiate between drift rates of saccade initiation and saccade generation.

Comparison of Other Factors to Free Viewing

Many dependencies of fixation duration are described in the literature and consequently modeled in our study. Comparing our results to the literature, it is obvious that our observed effects are drastically smaller, and sometimes not existent. One example is the absolute angle of a saccade: saccades from the center to above the horizon are generally observed to be faster than saccades below it (for an overview see Greene, Brown, & Dauphin, 2014). For example Heywood & Churcher (1980) reported an effect of 31 ms and Tzelepi, Yang and Kapoula, 2005 reported 27 ms. Our effect is estimated to be around 0.8 ms and thus does not exist in practical terms. Another example is the amplitude of the previous saccade, which has been described in a search paradigm by (Salthouse & Ellis, 1980) with ~90 ms for 15°. In our study, we found only 2.6 ms. A third example is an angle to the previous fixation, termed saccadic momentum or inhibition to return (Anderson, Yadav, & Carpenter, 2008; Dorris, Taylor, Klein, & Munoz, 1999; Wilming et al., 2013), which has been found to influence fixation duration with a piecewise linear relationship by approximately 70 ms per 180° (Wilming et al., 2013). In our study we did not find evidence for such an effect. Note that some effects do increase in size when not taking into account shared variance with other predictors. This is evident from the mismatch between effect size and marginal data as plotted in the figures. Still, even interpreting the marginal effects, the same qualitative judgment of smaller effects than in free viewing persists.

Some of the relationships described in the literature seem to be task dependent. For example, in a recent study (Nuthmann, 2016) a saccadic momentum effect could only be found for memorization and aesthetic preference but only very weak during a search task. This could partially explain the mismatch of our effect sizes. Yet another factor is that in our paradigm,

peripheral information is restricted and the controlled fixation duration might have a strong influence on the oculomotor system to show the same effects that can be seen in unrestricted viewing (Nuthmann, 2014). We propose a third, nonexclusive explanation: in our study, we separate the processing of the current input from the planning and execution of the next saccade. Also, our observed choice times are quite fast. In free viewing, these two processes occur in parallel. Thus, saccadic planning can start quite early in the current fixation. Consequently, in free viewing more time is available for other processes to intervene and modify either the duration of the exploitation phase or the planning of the next saccade. To test whether the exploitation phase is modified in free viewing by these geometric factors, we propose a modification to our experiment: instead of hiding the currently exploited bubble and stopping the exploitation process, we ought to keep it together with the new target bubbles. Thus, exploitation can continue and is partially under the subjects' control, and we expected to see effect sizes of similar size found in free viewing.

Comparison to Nuthmann (2016)

In a recent study by Nuthmann (2016), a complementary approach to this study was used. She analyzed fixation durations from free viewing with three different tasks using linear mixed models. Of course, the paradigm differs in many ways. Nuthmann used unrestricted free viewing with different tasks, but we used the combination of the bubble and guided-viewing paradigm, a restricted viewing task. Nuthmann focused on the influence of stimulus features on the fixation durations, we merely analyzed the difference between background images based on pink noise against urban images as our focus laid on the direct control of fixation durations. Both studies analyzed oculomotor effects but have a slightly different set of predictors. In this study, we additionally included the absolute angle and the absolute position of the fixation.

In general, Nuthmann found very strong oculomotor effects over all tasks, with some exceptions. For example, Nuthmann observed strong saccadic momentum in two tasks, but much weaker momentum in the visual search task. After controlling for other oculomotor effects, the initially observed distance-to-center effect of Nuthmann vanished. This was not the case in our study. We found an effect of distance to center even though modeling very similar other oculomotor predictors. Building upon the work of Nuthmann on feature influences, our paradigm could be used to directly check their saliency-based findings on a causal level: one can modify the low-level features of certain bubbles. For example, in a new experiment one can experimentally modify the "clutteredness" of certain bubbles and then expect a modified choice time.

Alternative Statistical Models, and Model Critique

We analyzed the choice times using a Bayesian linear mixed model. Linear mixed models (Bates, Mächler, Bolker, & Walker, 2014) and other hierarchical models (Gelman & Hill, 2007) are steadily replacing the need to use traditional ANOVA/ANCOVA (Baayen & Milin, 2010; Bagciella, Sloan, & Heitjan, 2000; Kliegl, Masson, & Richter, 2010; Nuthmann, 2016; Quené & Van Den Bergh, 2004; Richter, 2006). On the other hand, statistical recommendations for applied mixed models are still being developed (Barr, Levy, Scheepers, & Tily, 2013; Bates, Kliegl, Vasishth, & Baayen, 2015) and sometimes puzzling results can be observed (Hodges, 2014). One common problem is that maximum likelihood estimates fail to converge, mostly due to the complexity of the covariance matrix between random effects (Barr et al., 2013; Bates et al., 2015). Due to the high number of predictors our model is prone to this problem. Therefore, we used a Bayesian version of the mixed model (Sorensen et al., 2016) that allows us to put a small prior on the covariance matrix that de-emphasizes correlations between random slopes. This allowed us to fit the complex model and interpret the maximal instead of a reduced (possibly parsimonious) mixed model. In the case sufficient evidence from the data for a nonzero correlation exists, this will overrule this weak prior. This contrasts with forcing correlations to be zero as commonly done in more parsimonious models. It is currently up for debate whether it is possible to use those parsimonious models while still preserving correct type-1 error rates (Bates et al., 2015; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2015). Often only Bayesian mixed models allow estimating all random slopes in the mixed model to achieve a full statistical coverage when the model is sufficiently complex. Without all random slopes, posteriors, standard errors, or confidence intervals being too narrow, resulting p values are too small, type-1 errors are too high, and thus the effect estimates are too liberal (Barr et al., 2013; Schielzeth & Forstmeier, 2009; Sorensen et al., 2016). After the model fit, we used posterior predictive model checks to see whether the model adequately captures the data. As expected from a Gaussian fit of skewed data, the model checks revealed that we adequately fitted the mean of the distributions, but failed to model the tails appropriately. A possible enhancement for future analysis is to model the choice times using skewed normals, lognormals, or mixed normal distribution to reflect the most extreme quantiles.

The Bubbled Guided-Viewing Paradigm as a Research Tool

We introduced a new paradigm to investigate fixation durations in free viewing. There are some benefits over current free viewing paradigms, but also shortcomings. One problem with this paradigm is that we do not have a satisfying explanation of why the geometrical correlative effects are so much smaller than in free viewing. This could be a potential problem in studying any geometric effect with this paradigm. It might be that using a less naïve forced fixation duration distribution could alleviate this problem. Second, due to strict criteria, we had to remove ~45% of the data. The biggest factor was intermediate fixations between two bubbles.

Further analysis is needed to determine whether better online fixation detection algorithms or higher accuracy can improve upon this. On the other hand, the benefits are quite clear: we have a direct window to the decision-making mechanisms during free viewing. It controls the time a stimulus is available, how many and where the next stimuli are displayed, and thus, gives valuable variance estimates to constrain computational models. The paradigm is easily expandable to other spatial sampling mechanisms. In this case, we used a Gaussian distance measure, but it is straightforward to relate the geometrics of exploration to real free-viewing scan paths of subjects recorded beforehand. An alternative could be the use of empirical saliency maps that mean sampling the points as bubbles that are most likely fixated by other subjects. In addition, it is not clear how the number-of-bubbles effect generalizes to more bubbles, but there are two likely candidates: an asymptotic behavior or an inverse cubic function. Summarizing the benefits, we can state that we have a new tool to experimentally probe the exploration–exploitation state of the system that allows for close control and is flexible in its sampling scheme.

Conclusions

Here we developed a new paradigm that allows for experimental control over fixation durations and exploration behavior. We observed that we could selectively interrupt exploitation behavior and confirm predictions from the exploration–exploitation idea. In addition, we show a monotonic increase of choice time with the number of future targets.

Acknowledgements

The authors thank Jiameng Wu for helping with the data collection and analysis. This work was supported by SPP 1665 and the European Union through the H2020 FET Proactive Project socSMCs (GA no 641321).

References

- Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., ... Zuni, K. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), 253–267. <http://doi.org/10.1126/science.aac4716>
- Afsari, Z., Ossandón, J. P., & König, P. (2016). The dynamic effect of reading direction habit on spatial asymmetry of image perception. *Journal of Vision*, 16(11), 8. <http://doi.org/10.1167/16.11.8>
- Anderson, A. J., Yadav, H., & Carpenter, R. H. S. (2008). Directional Prediction by the Saccadic System. *Current Biology*, 18(8), 614–618. <http://doi.org/10.1016/j.cub.2008.03.057>
- Baayen, R. H., & Milin, P. (2010). Analyzing Reaction Times. *International Journal of Psychological Research*, 3(2), 12–28. <http://doi.org/10.1287/mksc.12.4.395>
- Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, 37(1), 13–20. <http://doi.org/10.1017/S0048577200980648>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <http://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D. M. (2010). lme4: Mixed-effects modeling with R, 131. <http://doi.org/10.1177/009286150103500418>
- Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv Preprint arXiv:1506.04967*, 1–27. <http://doi.org/arXiv:1506.04967>
- Bates, D. M., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. Available Online at: <http://CRAN.R-Project.org/package=lme4>. Retrieved from <http://arxiv.org/abs/1406.5823>
- Baumeister, A., & Joubert, C. E. (1969). Interactive effects on reaction time of preparatory interval length and preparatory interval frequency. *Journal of Experimental Psychology*, 82(2), 393–395. <http://doi.org/10.1037/h0028119>
- Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, 19(9), 967–983. [http://doi.org/10.1016/0042-6989\(79\)90222-0](http://doi.org/10.1016/0042-6989(79)90222-0)
- Beintema, J. A., van Loon, E. M., & van den Berg, A. V. (2005). Manipulating saccadic decision-rate distributions in visual search. *Journal of Vision*, 5(3), 1. <http://doi.org/10.1167/5.3.1>
- Berger-Tal, O., Nathan, J., Meron, E., & Saltz, D. (2014). The exploration-exploitation dilemma: A multidisciplinary framework. *PLoS ONE*, 9(4). <http://doi.org/10.1371/journal.pone.0095693>
- Bridson, R. (2007). Fast Poisson disk sampling in arbitrary dimensions. In *ACM SIGGRAPH*

- 2007 sketches on - SIGGRAPH '07 (p. 22–es). New York, New York, USA: ACM Press.
<http://doi.org/10.1145/1278780.1278807>
- Buswell, G. T. (1935). How people look at pictures: A study of the psychology of perception in art. *Chicago: University of Chicago Press*. <http://doi.org/citeulike-article-id:3847289>
- Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., & Durand, F. (2016). What do different evaluation metrics tell us about saliency models? <http://arxiv.org/abs/1604.03605>. Retrieved from <http://arxiv.org/abs/1604.03605>
- Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: a probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32. <http://doi.org/10.18637/jss.v076.i01>
- Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, 377(6544), 59–62. <http://doi.org/10.1038/377059a0>
- Carr, D., Lewin-Koh, N., & Maechler, M. (2015). hexbin: Hexagonal Binning Routines.
- Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience*, 11(7), 851–851. <http://doi.org/10.1038/nn0708-851c>
- Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 362(1481), 933–942. <http://doi.org/10.1098/rstb.2007.2098>
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45. <http://doi.org/no DOI found>
- Daw, N. D., O'Doherty, J. P., Dayan, P., Dolan, R. J., & Seymour, B. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095), 876–9. <http://doi.org/10.1038/nature04766>
- Dorris, M. C., Taylor, T. L., Klein, R. M., & Munoz, D. P. (1999). Influence of previous visual stimulus or saccade on saccadic reaction times in monkey. *Journal of Neurophysiology*, 81(5), 2429–2436.
- Drazin, D. H. (1961). Effects of foreperiod, foreperiod variability, and probability of stimulus occurrence on simple reaction time. *Journal of Experimental Psychology*, 62(1), 43–50. <http://doi.org/10.1037/h0046860>
- Einhäuser, W., & Nuthmann, A. (2016). Salient in space, salient in time: Fixation probability predicts fixation duration during natural scene viewing. *Journal of Vision*, 16(11), 13. <http://doi.org/10.1167/16.11.13>
- Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, 8(14), 18–18. <http://doi.org/10.1167/8.14.18>

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review*, 112(4), 777–813. <http://doi.org/10.1037/0033-295X.112.4.777>

Findlay, J. M., & Harris, L. R. (1984). Small Saccades to Double-Stepped Targets Moving in Two Dimensions. *Advances in Psychology*, 22(C), 71–78. [http://doi.org/10.1016/S0166-4115\(08\)61820-8](http://doi.org/10.1016/S0166-4115(08)61820-8)

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behav Brain Sci*, 22, 661–721. <http://doi.org/10.1017/S0140525X99002150>

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, 8(2), 1–17. <http://doi.org/10.1167/8.2.6>

Foulsham, T., & Underwood, G. (2009). Does conspicuity enhance distraction? Saliency and eye landing position when searching for objects. *Quarterly Journal of Experimental Psychology* (2006), 62(6), 1088–1098. <http://doi.org/10.1080/17470210802602433>

Gameiro, R. R., Kaspar, K., König, S. U., Nordholt, S., & König, P. (2017). Exploration and Exploitation in Natural Viewing Behavior. *Scientific Reports*, 7(1), 2311. <http://doi.org/10.1038/s41598-017-02526-1>

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition*.

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. *Policy Analysis*, 1–651. <http://doi.org/10.2277/0521867061>

Glaholt, M., Rayner, K., & Reingold, E. (2012). The mask-onset delay paradigm and the availability of central and peripheral visual information during scene viewing. *Journal of Vision*, 12(1), 9–9. <http://doi.org/10.1167/12.1.9.Introduction>

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, 30(1), 535–574. <http://doi.org/10.1146/annurev.neuro.29.051605.113038>

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, 41(17), 2261–2271. [http://doi.org/10.1016/S0042-6989\(01\)00097-9](http://doi.org/10.1016/S0042-6989(01)00097-9)

Gould, J. D. (1973). Eye movements during visual search and memory search. *Journal of Experimental Psychology*, 98(1), 184–95. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3997357&tool=pmcentrez&rendertype=abstract>

Greene, H. H., Brown, J. M., & Dauphin, B. (2014). When do you look where you look? A visual field asymmetry. *Vision Research*, 102, 33–40. <http://doi.org/10.1016/j.visres.2014.07.012>

1019 Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive*
1020 *Sciences*, 9(4), 188–194.

1021 Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general
1022 mechanism for perceptual decision-making in the human brain. *Nature*, 431(7010),
1023 859–862. <http://doi.org/10.1038/nature02966>

1024 Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate
1025 human perceptual decision making. *Nature Reviews Neuroscience*, 9(6), 467–479.
1026 <http://doi.org/10.1038/nrn2374>

1027 Henderson, J. M., Olejarczyk, J., Luke, S. G., & Schmidt, J. (2014). Eye movement control
1028 during scene viewing: Immediate degradation and enhancement effects of spatial
1029 frequency filtering. *Visual Cognition*, 22(3–4), 486–502.
1030 <http://doi.org/10.1080/13506285.2014.897662>

1031 Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: evidence
1032 for mixed control of fixation durations. *Psychonomic Bulletin & Review*, 15(3), 566–573.
1033 <http://doi.org/10.3758/PBR.15.3.566>

1034 Henderson, J. M., & Smith, T. J. (2009). How are eye fixation durations controlled during
1035 scene viewing? Further evidence from a scene onset delay paradigm. *Visual Cognition*,
1036 17(February 2015), 1055–1082. <http://doi.org/10.1080/13506280802685552>

1037 Heywood, S., & Churcher, J. (1980). Structure of the visual array and saccadic latency:
1038 implications for oculomotor control. *The Quarterly Journal of Experimental Psychology*,
1039 32(2), 335–341. <http://doi.org/10.1080/14640748008401169>

1040 Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental*
1041 *Psychology*, 4(1), 11–26. <http://doi.org/10.1080/17470215208416600>

1042 Hodges, J. S. (2014). *Richly Parameterized Linear Models: Additive, Time Series, and*
1043 *Spatial Models Using Random Effects*. Florida: Chapman and Hall. Retrieved from
1044 [https://books.google.de/books?hl=de&lr=&id=3VL8AQAAQBAJ&oi=fnd&pg=PP1&dq=2014+hodges+Richly+Parameterized+Linear+Models&ots=edWQk1hf-](https://books.google.de/books?hl=de&lr=&id=3VL8AQAAQBAJ&oi=fnd&pg=PP1&dq=2014+hodges+Richly+Parameterized+Linear+Models&ots=edWQk1hf-T&sig=6og6sil2R8WU1AOZqDnAyENzVVM#v=onepage&q=2014+hodges+Richly+Parameterized+Linear+Models&f=false)
1045 [T&sig=6og6sil2R8WU1AOZqDnAyENzVVM#v=onepage&q=2014+hodges+Richly](https://books.google.de/books?hl=de&lr=&id=3VL8AQAAQBAJ&oi=fnd&pg=PP1&dq=2014+hodges+Richly+Parameterized+Linear+Models&ots=edWQk1hf-T&sig=6og6sil2R8WU1AOZqDnAyENzVVM#v=onepage&q=2014+hodges+Richly+Parameterized+Linear+Models&f=false)
1046 [Parameterized Linear Models&f=false](https://books.google.de/books?hl=de&lr=&id=3VL8AQAAQBAJ&oi=fnd&pg=PP1&dq=2014+hodges+Richly+Parameterized+Linear+Models&ots=edWQk1hf-T&sig=6og6sil2R8WU1AOZqDnAyENzVVM#v=onepage&q=2014+hodges+Richly+Parameterized+Linear+Models&f=false)

1047

1048 Homan, M., & Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in
1049 Hamiltonian Monte Carlo. *The Journal of Machine Learning Research*. Retrieved from
1050 <http://dl.acm.org/citation.cfm?id=2638586>

1051 Hooge, I. T. C., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during
1052 visual search. *Vision Research*, 39(8), 1567–1575. [http://doi.org/10.1016/S0042-](http://doi.org/10.1016/S0042-6989(98)00213-2)
1053 [6989\(98\)00213-2](http://doi.org/10.1016/S0042-6989(98)00213-2)

1054 Hooge, I. T. C., Over, E. A. B., Van Wezel, R. J. A., & Frens, M. A. (2005). Inhibition of
1055 return is not a foraging facilitator in saccadic search and free viewing. *Vision Research*,

45(14), 1901–1908. <http://doi.org/10.1016/j.visres.2005.01.030>

Itti, L., Koch, C., & Niebur, E. (1998). A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254–1259.

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), 26–26. <http://doi.org/10.1167/11.13.26>

Kietzmann, T. C., & König, P. (2015). Effects of contextual information and stimulus ambiguity on overt visual sampling behavior. *Vision Research*, 110(Pt A), 76–86. <http://doi.org/10.1016/j.visres.2015.02.023>

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, 18(5), 655–681. <http://doi.org/10.1080/13506280902986058>

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219–27. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3836989>

Koehler, K., Guo, F., Zhang, S., & Eckstein, M. P. (2014). What do saliency models predict? *Journal of Vision*, 14(3), 14.

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>

König, P., Wilming, N., Kietzmann, T. C., Ossandon, J. P., Onat, S., Ehinger, B., ... Kaspar, K. (2016). Eye movements as a window to cognitive processes. *Journal of Eye Movement Research*, 9(3), 1–16. <http://doi.org/http://dx.doi.org/10.16910/jemr.9.5.3>

Kümmerer, M., Wallis, T. S. A., & Bethge, M. (2015). Information-theoretic model comparison unifies saliency metrics. *Proceedings of the National Academy of Sciences of the United States of America*, 112(52), 16054–9. <http://doi.org/10.1073/pnas.1510393112>

Lawrence, B. M. (2010). An anti-Hick's effect for exogenous, but not endogenous, saccadic eye movements. *Experimental Brain Research*, 204, 115–118. <http://doi.org/10.1007/s00221-010-2301-8>

Lawrence, B. M., St John, A., Abrams, R. a, & Snyder, L. H. (2008). An anti-Hick's effect in monkey and human saccade reaction times. *Journal of Vision*, 8(3), 26.1-7. <http://doi.org/10.1167/8.3.26>

Leach, J. C. D., & Carpenter, R. H. S. (2001). Saccadic choice with asynchronous targets: evidence for independent randomisation. *Vision Research*, 41(25–26), 3437–3445.

1093 [http://doi.org/10.1016/S0042-6989\(01\)00059-1](http://doi.org/10.1016/S0042-6989(01)00059-1)

1094 Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use
 1095 standard deviation around the mean, use absolute deviation around the median.
 1096 *Journal of Experimental Social Psychology*, 49(4), 764–766.
 1097 <http://doi.org/10.1016/j.jesp.2013.03.013>

1098 Loschky, L., McConkie, G., Yang, J., & Miller, M. (2005). The limits of visual resolution in
 1099 natural scene viewing. *Visual Cognition*, 12(6), 1057–1092.
 1100 <http://doi.org/10.1080/13506280444000652>

1101 Ludwig, C. J. H. (2009). Temporal integration of sensory evidence for saccade target
 1102 selection. *Vision Research*, 49(23), 2764–2773.
 1103 <http://doi.org/10.1016/j.visres.2009.08.012>

1104 Ludwig, C. J. H., Davies, J. R., & Eckstein, M. P. (2014). Foveal analysis and peripheral
 1105 selection during active visual sampling. *Proceedings of the National Academy of*
 1106 *Sciences of the United States of America*, 111(2).
 1107 <http://doi.org/10.1073/pnas.1313553111>

1108 Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. M. (2015). Balancing Type I
 1109 Error and Power in Linear Mixed Models. *arXiv*, (2013), 1–14.
 1110 <http://doi.org/10.1016/j.jml.2017.01.001>

1111 Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual
 1112 search. *Vision Research*, 38(12), 1805–1815. [http://doi.org/10.1016/S0042-](http://doi.org/10.1016/S0042-6989(97)00349-0)
 1113 [6989\(97\)00349-0](http://doi.org/10.1016/S0042-6989(97)00349-0)

1114 Mowbray, G. H. (1964). Subjective expectancy and choice reaction times. *Quarterly Journal*
 1115 *of Experimental Psychology*, 16, 216-223 .

1116 Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search.
 1117 *Nature*, 434(7031), 387–91. <http://doi.org/10.1038/nature03390>

1118 Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological*
 1119 *Bulletin*, 89(1), 133. <http://doi.org/10.1037/0033-2909.89.1.133>

1120 Nuthmann, A. (2014). How do the regions of the visual field contribute to object search in
 1121 real-world scenes? Evidence from eye movements. *Journal of Experimental*
 1122 *Psychology: Human Perception and Performance*, 40(1), 342–360.
 1123 <http://doi.org/10.1037/a0033854>

1124 Nuthmann, A. (2016). Fixation durations in scene viewing: Modeling the effects of local
 1125 image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review*, 1–
 1126 23. <http://doi.org/10.3758/s13423-016-1124-4>

1127 Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational
 1128 model of fixation durations in scene viewing. *Psychological Review*, 117(2), 382–405.
 1129 <http://doi.org/10.1037/a0018924>

1130 Ossandón, J. P., Onat, S., Cazzoli, D., Nyffeler, T., Müri, R., & König, P. (2012). Unmasking
 1131 the contribution of low-level features to the guidance of attention. *Neuropsychologia*,
 1132 50(14), 3478–3487. <http://doi.org/10.1016/j.neuropsychologia.2012.09.043>
 1133 Ossandón, J. P., Onat, S., & König, P. (2014). Spatial biases in viewing behavior. *Journal of*
 1134 *Vision*, 14(2), 20–20. <http://doi.org/10.1167/14.2.20>
 1135 Pannasch, S., Schulz, J., & Velichkovsky, B. M. (2011). On the control of visual fixation
 1136 durations in free viewing of complex images. *Attention, Perception & Psychophysics*,
 1137 73(4), 1120–32. <http://doi.org/10.3758/s13414-011-0090-1>
 1138 Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on
 1139 replicability in psychological science: A crisis of confidence? *Perspectives on*
 1140 *Psychological Science*, 7(6), 528–530. <http://doi.org/10.1177/1745691612465253>
 1141 Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R-Core-Team. (2016). {nlme}: Linear and
 1142 Nonlinear Mixed Effects Models.
 1143 Posner, M. I., & Cohen, Y. (1980). Attention and the control of movements. In *Tutorials in*
 1144 *motor behavior* (pp. 243–258).
 1145 Potter, M. C. (1976). Short-Term Conceptual Memory for Pictures. *Journal of Experimental*
 1146 *Psychology: Human Learning and Memory*, 2(5), 509–522. [http://doi.org/10.1037/0278-](http://doi.org/10.1037/0278-7393.2.5.509)
 1147 7393.2.5.509
 1148 Proctor, R. W., & Schneider, D. W. (2017). Hick's Law for Choice Reaction Time: A Review.
 1149 *The Quarterly Journal of Experimental Psychology*, 218(April), 1–56.
 1150 <http://doi.org/10.1080/17470218.2017.1322622>
 1151 Quené, H., & Van Den Bergh, H. (2004). On multi-level modeling of data from repeated
 1152 measures designs: A tutorial. *Speech Communication*, 43(1–2), 103–121.
 1153 <http://doi.org/10.1016/j.specom.2004.02.004>
 1154 R Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna,
 1155 Austria. Retrieved from <http://www.r-project.org/>
 1156 Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*.
 1157 <http://doi.org/10.1037/0033-2909.114.3.510>
 1158 Ratcliff, R. (2001). Diffusion and random walk processes. *International Encyclopedia of the*
 1159 *Social and Behavioral Sciences*, 6(1977), 3668–3673. Retrieved from
 1160 <http://www.sciencedirect.com/science/article/pii/B0080430767006203>
 1161 Rayner, K. (1978). Eye movements in reading and information processing. *Psychological*
 1162 *Bulletin*, 85(3), 618–660. <http://doi.org/10.1037/0033-2909.85.3.618>
 1163 Rayner, K. (1998). Eye Movements in Reading and Information Processing : 20 Years of
 1164 Research, 124(3), 372–422.
 1165 Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual
 1166 search. *The Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506.

1167 <http://doi.org/10.1080/17470210902816461>

1168 Rayner, K., Inhoff, a W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981).
 1169 Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of*
 1170 *Experimental Psychology. Human Perception and Performance*, 7(1), 167–179.
 1171 <http://doi.org/10.1037/0096-1523.7.1.167>

1172 Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2009). Eye movements and
 1173 visual encoding during scene perception. *Psychological Science*, 20(1), 6–10.
 1174 <http://doi.org/10.1111/j.1467-9280.2008.02243.x>

1175 Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye
 1176 movement control in reading. *Psychological Review*, 105(1), 125–157.
 1177 <http://doi.org/10.1037/0033-295X.105.1.125>

1178 Reingold, E. M., & Stampe, D. M. (2002). Saccadic inhibition in voluntary and reflexive
 1179 saccades. *Journal of Cognitive Neuroscience*, 14(3), 371–388.
 1180 <http://doi.org/10.1162/089892902317361903>

1181 Richter, T. (2006). What Is Wrong With ANOVA and Multiple Regression? Analyzing
 1182 Sentence Reading Times With Hierarchical Linear Models. *Discourse Processes*, 41(3),
 1183 221–250. http://doi.org/10.1207/s15326950dp4103_1

1184 Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where
 1185 you look. *Journal of Vision*, 7(14), 16–20.

1186 Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for
 1187 estimating response time distributions. *Psychonomic Bulletin & Review*, 12(2), 195–
 1188 223. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16082801>

1189 Salthouse, T. A., & Ellis, C. L. (1980). Determinants of eye-fixation duration. *The American*
 1190 *Journal of Psychology*. <http://doi.org/10.2307/1422228>

1191 Schall, J. D., & Thompson, K. G. (1999). neural selection and control of visually guided eye
 1192 movements. *Annual Review of Neuroscience*, 22(1), 241–259.
 1193 <http://doi.org/10.1146/annurev.neuro.22.1.241>

1194 Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: overconfident
 1195 estimates in mixed models. *Behavioral Ecology : Official Journal of the International*
 1196 *Society for Behavioral Ecology*, 20(2), 416–420. <http://doi.org/10.1093/beheco/arn145>

1197 Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous
 1198 neural activity prior to self-initiated movement. *Proceedings of the National Academy of*
 1199 *Sciences of the United States of America*, 109(42), E2904-13.
 1200 <http://doi.org/10.1073/pnas.1210467109>

1201 Shadlen, M. N., Britten, K., Newsome, W., & Movshon, J. (1996). A computational analysis
 1202 of the relationship between neuronal and behavioral responses to visual motion. *J.*
 1203 *Neurosci.*, 16(4), 1486–1510. Retrieved from

1204 <http://www.jneurosci.org/content/16/4/1486.short>
 1205 Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the
 1206 parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol*, 86(4), 1916–1936.
 1207 Retrieved from <http://jn.physiology.org/content/86/4/1916.short>
 1208 Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends*
 1209 *in Neurosciences*, 27(3), 161–168. <http://doi.org/10.1016/j.tins.2004.01.006>
 1210 Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian Linear Mixed Models using
 1211 Stan: A tutorial for psychologists, linguists, and cognitive scientists. *The Quantitative*
 1212 *Methods for Psychology*, 12(3), 175–200. Methodology. Retrieved from
 1213 <http://arxiv.org/abs/1506.06201>
 1214 Tatler, B. W., Baddeley, R. J., & Vincent, B. T. (2006). The long and the short of it: Spatial
 1215 statistics at fixation vary with saccade amplitude and task. *Vision Research*, 46(12),
 1216 1857–1862. <http://doi.org/10.1016/j.visres.2005.12.005>
 1217 Tatler, B. W., & Vincent, B. T. (2009). The prominence of behavioural biases in eye
 1218 guidance. *Visual Cognition*, 17(6–7), 1029–1054.
 1219 <http://doi.org/10.1080/13506280902764539>
 1220 Trukenbrod, H. a, & Engbert, R. (2014). ICAT: a computational model for the adaptive
 1221 control of fixation durations. *Psychonomic Bulletin & Review*, 21(4), 907–34.
 1222 <http://doi.org/10.3758/s13423-013-0575-0>
 1223 Tzelepi, A., Yang, Q., & Kapoula, Z. (2005). The effect of transcranial magnetic stimulation
 1224 on the latencies of vertical saccades. *Experimental Brain Research*, 164(1), 67–77.
 1225 <http://doi.org/10.1007/s00221-005-2250-9>
 1226 Van Den Berg, A. V., & Van Loon, E. M. (2005). An invariant for timing of saccades during
 1227 visual search. *Vision Research*, 45(12), 1543–1555.
 1228 <http://doi.org/10.1016/j.visres.2004.12.018>
 1229 van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy Array: A Structure for
 1230 Efficient Numerical Computation. *Computing in Science & Engineering*, 13(2), 22–30.
 1231 <http://doi.org/10.1109/MCSE.2011.37>
 1232 VanRullen, R., Reddy, L., & Koch, C. (2004). Visual Search and Dual Tasks Reveal Two
 1233 Distinct Attentional Resources. *Journal of Cognitive Neuroscience*, 16(1), 4–14.
 1234 <http://doi.org/10.1162/089892904322755502>
 1235 Vaughan, J. (1982). Control of fixation duration in visual search and memory search:
 1236 Another look. *Journal of Experimental Psychology: Human Perception and*
 1237 *Performance*, 8(5), 709–723. <http://doi.org/10.1037/0096-1523.8.5.709>
 1238 Vaughan, J., & Graefe, T. M. (1977). Delay of stimulus presentation after the saccade in
 1239 visual search, 22(2), 201–205.
 1240 Vickers, D. (1970). Evidence for an Accumulator Model of Psychophysical Discrimination.

Ergonomics, 13(1), 37–58. <http://doi.org/10.1080/00140137008931117>

Walshe, R. C., & Nuthmann, A. (2014). Asymmetrical control of fixation durations in scene viewing. *Vision Research*, 100, 38–46. <http://doi.org/10.1016/j.visres.2014.03.012>

Walshe, R. C., & Nuthmann, A. (2015). Mechanisms of saccadic decision making while encoding naturalistic scenes. *Journal of Vision*, 15, 1–38. <http://doi.org/10.1167/15.5.21>

White, B. J., Berg, D. J., Kan, J. Y., Marino, R. A., Itti, L., & Munoz, D. P. (2017). Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. *Nature Communications*, 8, 14263. <http://doi.org/10.1038/ncomms14263>

Wickham, H. (2007). Reshaping Data with the {reshape} Package. *Journal of Statistical Software*, 21(12), 1--20.

Wickham, H. (2015). stringr: Simple, Consistent Wrappers for Common String Operations.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behav Res Methods*, 42(3), 671–684. <http://doi.org/10.3758/brm.42.3.671>

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>

Wilming, N., Onat, S., Ossandón, J. P., Açık, A., Kietzmann, T. C., Kaspar, K., ... König, P. (2017). An extensive dataset of eye movements during viewing of complex images. *Scientific Data*, 4, 160126. <http://doi.org/10.1038/sdata.2016.126>

Wu, C.-C., & Kowler, E. (2013). Timing of saccadic eye movements during visual search for multiple targets. *Journal of Vision*, 13(11), 1–21. <http://doi.org/10.1167/13.11.11>