Large Pupils Predict Goal-driven Eye Movements

Sebastiaan Mathôt[1], Alisha Siebold[2,3], Mieke Donk[2], and Françoise Vitu[1]

[1] Aix-Marseille University, CNRS, LPC UMR 7290, Marseille, France

[2] Dept. of Cognitive Psychology, VU University Amsterdam

[3] Center for Mind/Brain Sciences, University of Trento, Italy

Address for correspondence:

Aix-Marseille University, CNRS

Laboratoire de Psychologie Cognitive, UMR 7290

3 Place Victor Hugo

Centre St. Charles, Bâtiment 9, Case D

13331 Marseille

France

Author note:

This article was accepted for publication in the *Journal of Experimental Psychology: General* on March 12, 2015. This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Abstract

Here we report that large pupils predict fixations of the eye on low-salient, inconspicuous parts of a visual scene. We interpret this as showing that mental effort, reflected by a dilation of the pupil, is required to guide gaze towards objects that are relevant to current goals, but may not be very salient. When mental effort is low, reflected by a constriction of the pupil, the eyes tend to be captured by high-salient parts of the image, irrespective of top-down goals. The relationship between pupil size and visual saliency was not driven by luminance, nor a range of other factors that we considered. Crucially, the relationship was strongest when mental effort was invested exclusively in eye-movement control (i.e. reduced in a dual-task setting), which suggests that it is not due to general effort or arousal. Our finding illustrates that goal-driven control during scene viewing requires mental effort, and that pupil size can be used as an on-line measure to track the goal-drivenness of behavior.

Large Pupils Predict Goal-driven Eye Movements

Behavior is guided to a considerable extent by the environment. Imagine that you are walking on a beach, when an approaching object grabs your attention (Franconeri & Simons, 2003). You instinctively duck. Only later do you realize that the approaching object was a volleyball, and that your evasive maneuver prevented a collision between the volleyball and your head. In situations such as these, behavior is based directly on salient properties of the environment (movement in this case) that automatically attract attention and elicit a response. This type of bottom-up control allows you to respond quickly and adaptively to dangers and opportunities in the environment.

But bottom-up control does not address all challenges of the environment equally well. Some situations require an in-depth, top-down analysis. For example, during driving you need to pay attention to things that are not salient, but nevertheless relevant (rear-view mirror, speed indicator, pedestrians, etc.; Recarte & Nunes, 2003). Radiologists face a similar challenge when they try to identify rare and low-salient patterns in medical images (Wolfe, 2010). In neither of these situations does purely bottom-up control lead to the most adaptive behavior. Therefore, it is important to understand the relationship between bottom-up and top-down control. Which factors determine whether top-down or bottom-up control is dominant in a particular situation? And is it possible to track the goal-drivenness of behavior as it unfolds?

Van Zoest, Donk, and Theeuwes (2004; see also Donk & van Zoest, 2008; Siebold, van Zoest, & Donk, 2011; van Zoest & Donk, 2005, 2006, 2008) addressed these questions by considering the time-course of visual selection. In one experiment (Exp. 4; Van Zoest et al., 2004), participants searched for a pre-specified target, which was a line-segment tilted 45° from a vertical orientation. The target was embedded in a field of homogeneous vertical line-segments. In addition, a distractor was presented, which was also a tilted line-segment. Crucially, the distractor was either more salient than the target, equally salient, or less salient. The results were clear-cut. Short-latency eye movements (with a latency of less than 250 ms) overwhelmingly landed on the most salient object, regardless of whether it was the target or the distractor. In contrast, long-latency eye movements were mostly directed at the target stimulus, regardless of the saliency of the distractor. This suggests that short-latency eye movements are driven primarily by bottom-up saliency, whereas long-latency eye movements are driven primarily by top-down goals.

The findings of Van Zoest and colleagues (2004; Donk & van Zoest, 2008; Siebold et al., 2011; van Zoest & Donk, 2005, 2006, 2008; see also Godijn & Theeuwes, 2002) suggest that bottom-up control is dominant in guiding behavior 'early on' (i.e. for short-latency eye movements and/ or immediately after a stimulus has appeared), but that top-down control becomes dominant later in time (Foulsham & Underwood, 2007; Henderson, Malcolm, & Schandl, 2009; Parkhurst, Law, & Niebur, 2002). What we propose here is that mental effort is required to exert top-down control. We use the term 'mental effort' in connection with the pupillometry literature (cf. Beatty, 1982), but what we suggest specifically is the following. It is computationally non-trivial to evaluate the environment in terms of top-down goals, because goals are abstract concepts that cannot be mapped directly onto simple features (except for relatively simple goals, such as "looking for something red", cf. Wolfe, 1994). Presumably, this process takes time (cf. Van Zoest et al., 2004) and requires feedback loops between the early visual system and higher-order parietal and frontal areas (Lamme & Roelfsema, 2000). Moreover, top-down control requires inhibition of reflexive behavior. In this sense, mental effort is required, even though the computations that underlie top-down guidance may be wholly or partly implicit. Crucially, the amount of effort that we invest in a task fluctuates over time (Reimer et al., 2014). For example, participants in a search-for-Waldo experiment (e.g., R. M. Klein & MacInnes, 1999) may initially invest a lot of effort in trying to find Waldo. But over the course of many trials drowsiness sets in and effort wanes. At this point, participants let their eyes glaze over the image, from one conspicuous element to the next, still making eye movements, but no longer engaging the top-down goal to find Waldo.

Mental effort, as described above, is easily measured. As numerous studies have shown, the pupil faithfully dilates when you engage in a task that in some way requires mental effort (for recent reviews, see Laeng, Sirois, & Gredebäck, 2012; Sirois & Brisson, 2014; for a classic, but remarkably insightful review, see Loewenfeld, 1958): The seminal studies by Kahneman and Beatty (1966) have shown that the pupil dilates as a result of increased working memory load; Pupil dilation indicates the amount of effort invested in speech comprehension (Zekveld, Kramer, & Festen, 2010); More recently, and more directly related to the present study, it has been shown that pupillary dilation accompanies the detection of a target in a Rapid Serial Visual Presentation (RSVP) paradigm (Privitera, Renninger, Carney, Klein, & Aguilar, 2010; see also Wolff, Scholz, Akyurek, & Rijn, 2014), and that pupil-size indicates search difficulty in visual search (Porter, Troscianko, & Gilchrist, 2007) and multiple-object tracking (Alnaes et al., 2014).

In the present experiments, we measured pupil size while participants viewed images and made unconstrained eye movements. We consistently found, with various types of stimuli and different task instructions, that the pupil was relatively dilated, reflecting increased mental effort (cf. Beatty, 1982; Kahneman & Beatty, 1966; Zekveld et al., 2010), when participants fixated on low-salient locations. Crucially, the relationship between pupil size and fixation saliency was reduced when mental effort was invested in a secondary task, suggesting that it was not due to general arousal. We interpret our findings as showing that top-down control of eye movements relies on mental effort. Without mental effort, or if effort is invested elsewhere, the eyes are captured by high-salient objects, regardless of whether these are relevant to current goals.

Experiments 1 and 2

Method

Materials and availability

All experimental materials, where possible given license restrictions, are available from https://github.com/smathot/materials_for_P0010.5.

Datasets, participants, procedure, and apparatus

We analyzed data from two previously unpublished experiments that were conducted for a different purpose. Data were independently collected and contained a mixture of different trial types. From Exp. 1 we analyzed only trials without any visible manipulation to the display (16 participants; 100 trials per participant). Exp. 1 was a difficult visual search task, in which participants searched for a small letter (‘Z’ or ‘H’; 0.4°x0.4°) embedded in a natural scene. From Exp. 2, we analyzed all trials, which were a mixture of three task instructions (16 participants; 100 trials per participant; instruction varied between participants): a visual-search task, as in Exp. 1; a memory task, in which participants were (falsely) informed that they would be asked questions about the images; and a free-viewing task, in which no instructions were given. Eye movements were unconstrained.

Although the experiments were conducted at different locations, a similar experimental set-up was used. Eye movements were recorded monocularly with an EyeLink 1000, a video-based eye tracker sampling at 1000 Hz (SR Research, Mississauga, ON, Canada). Stimulus presentation was controlled by OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) using the PsychoPy back-end (Peirce, 2007). Observers participated in the experiments for course credit or money, reported normal or corrected vision, and signed a written consent form. The experiments were conducted with approval of the Ethics Board of the Faculty of Psychology and Education (VCWE; Exp. 1) and the local ethics committee of Aix-Marseille Université (Exp. 2).

Stimuli, saliency maps, and pupillary luminance maps

In Exp. 1, stimuli were 200 photographs of natural scenes from the UPenn natural image database (Tkačik et al., 2011). Of these 200 stimuli, 100 were randomly selected for each participant. In Exp. 2, stimuli were 50 photographs of natural scenes from the Campus Scene collection (Burge & Geisler, 2011), and 50 3D fractals generated with the program Mandelbulber (Marczak, 2012).

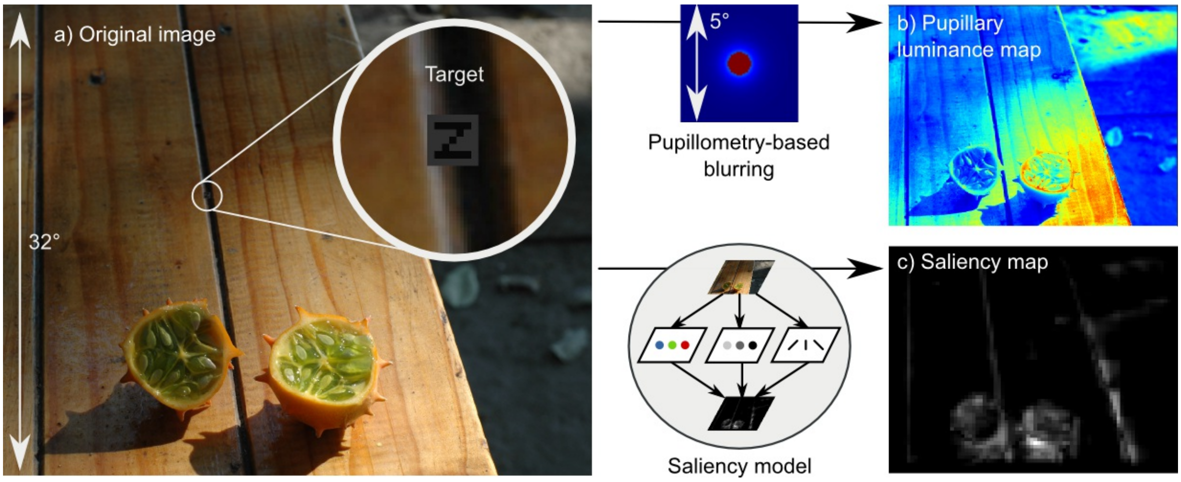

Figure 1. *a) Participants freely viewed images of natural scenes or three-dimensional fractals. A small 'Z' or 'H' was hidden in each image and served as the target on visual-search trials. b) For each image, a pupillary luminance map was created to predict the pupillary light response for each location. c) A saliency map was created to predict where the eyes would land if gaze was guided purely by bottom-up visual saliency.*

Saliency maps

For each image, we generated a saliency map, which is a best estimate of where the eyes would go if eye-movement guidance was purely bottom up (Figure 1c). Saliency maps were generated with the NeuroMorphic vision toolkit (Itti, Koch, & Niebur, 1998), using the 'Surprise' visual-cortex model. Saliency maps were obtained with the following command:

```
ezvision --just-initial-saliency-map --in=[input_filename] --out=png --vc-type=Surp --maxnorm-type=Surprise
```

Pupillary luminance maps

For each image, we generated a pupillary luminance map, which is a best effort to predict pupil size for each fixation if pupil size was purely determined by luminance (Figure 1b). In other words, for each pixel in the image we estimated a value (in arbitrary units) that would predict pupil size during fixations on that pixel if only luminance were a factor.

It is clear that the pupillary light response (PLR) is driven primarily by foveal illumination, with an exponential drop-off in sensitivity with increasing eccentricity (Crawford, 1936). Based on data from a pupillometry study by Hong, Narkiewicz, & Kardon (2001), we estimated the following relationship: $s = 33.2+10.6 \cdot e^{-11.2*ecc}$

Here *s* is pupil sensitivity (dB; a measure of the pupillary response to the light stimulus) and *ecc* is the eccentricity (°) of the light stimulus. However, Hong et al. (2001) probed a limited number of widely spaced locations and it is possible, if not likely, that the exponential drop-off in pupillary sensitivity that is observed in the wider visual field does not hold for the fovea. There is, to the best of our knowledge, no detailed mapping of pupillary sensitivity for the foveal area. We conducted a number of pilot experiments ourselves, which confirmed the general pattern observed by Hong et al. (2001) and Crawford (1936), but we failed to obtain a sufficiently clear foveal mapping to answer this question with any degree of certainty. Therefore, we applied Occam's razor and modeled the fovea as an area of 1° diameter with a uniform pupillary sensitivity. Non-foveal areas were modeled with the formula described above. Using this 'mixed luminance kernel', we first gray-scaled and then blurred the photos, thus obtaining pupillary luminance maps.

Analysis

Model selection

Like any set of data with unconstrained eye movements, our datasets contain many correlations that need to be taken into account when interpreting a relationship of interest, in our case between pupil size and the saliency of fixated locations (from now on: fixation saliency). For example, if high-salient locations would be brighter than low-salient locations, a light response would manifest itself as a pupillary constriction when fixating high-salient locations. (This is merely an example, see Supplementary Methods for the actual relationship between saliency and luminance.)

To deal with these correlations, we focus on the 'partial' relationship between pupil size and fixation saliency, i.e. the slope of the effect in a linear-mixed effects (LME) model that accounts for several other factors that might co-vary with pupil size. We used a combination of a data-driven and a confirmatory approach, in which we selected control predictors in a data-driven way, and used confirmatory testing for the effects of interest (see Barr, Levy, Scheepers, & Tily, 2013 for guidelines and a discussion).

More specifically, we started with a basic model with fixation saliency as dependent variable and a by-participant random intercept. Next, we progressively added the following control predictors as fixed effects: trial number, fixation number in trial, luminance of the fixated position (as read from the pupillary luminance map), eccentricity (distance between fixated position and display center), horizontal fixaton position, vertical fixation position, fixation duration, and size of following saccade. Our choice of control predictors was not motivated by any theoretical interest, but served to account for as many factors as possible that might co-vary with pupil size. We did not include control predictors that correlated strongly with other control predictors, such as properties of preceding and following fixations (but see Temporal Characteristics). After adding each control predictor, we tested whether the addition improved the quality of the model, by determining X^2 and the associated p-value from the log-likelihood of the two models (Baayen, Davidson, & Bates, 2008). If the predictor significantly ($p < .05$) improved the model, it was included in the model, otherwise it was discarded.

Finally, after having constructed a model in this way, we added pupil size as fixed effect, as well as by-participant random slopes for pupil size. (More generally, we used random slopes only for predictors of theoretical interest, and not for control predictors, cf. Barr et al., 2013). Fixed effects were considered reliable when $t > 2$ (cf. Baayen, 2008), although we will emphasize general patterns over significance of individual analyses. In the main text, we will describe only the effects of theoretical interest, but full LME models are listed in the Supplementary Methods.

Pupil-size transformation

The power of an analysis can be increased (or reduced) by transforming variables. For example, when analyzing response times (RTs), statistical power often increases when a logarithmic or inverse transformation is applied to the RTs (Ratcliff, 1993). To our knowledge, the effect of transforming pupil-size measures has not been investigated. Most researchers use area (e.g., Mathôt, Dalmaijer, Grainger, & Van der Stigchel, 2014; Mathôt, van der Linden, Grainger, & Vitu, 2013) or diameter (e.g., Binda, Pereverzeva, & Murray, 2013) as dependent measure. Presumably (or certainly, for our own work) the chosen measure is determined by the default units of the eye tracker.

We used the data from Exp. 1 to determine the optimal pupil-size transformation. First, we converted pupil size to *Z* scores, so that pupil-size measures were comparable between experiments (the EyeLink provides uncalibrated pupil-size measures that depend on the specifics of the set-up). Next, we constructed LME models as described above, adding pupil size as fixed effect after applying one of the following transformations: D (= diameter), D^2 (= area), D^3, $D^{0.5}$, D^{-1}, and log(D). This resulted in six models that differed only in the pupil-size transformation that was applied. Based on each model's log-likelihood, we selected the optimal transformation, i.e. from the model with the highest log-likelihood. This was the inverse transformation: D^{-1}. Therefore, all subsequent analyses were performed after applying an inverse transformation to *Z* scored pupil-size diameter. (For details, see Supplementary Methods.)

Results

Trial duration and number of fixations

Trials lasted on average 12.6 s (Exp. 1) and 14.6 s (Exp. 2). The average fixation durations were 285 ms (Exp. 1) and 294 ms (Exp. 2). A total of 59,585 (Exp. 1) and 64,526 (Exp. 2) fixations were entered into the analyses. No participants or data points were excluded.

Pupil size as a predictor of fixation saliency

The main result is that pupil size is related to fixation saliency: When the pupil is large, the average saliency of fixated locations is lower than when the pupil is small (Figure 2). This effect is robust, present for the duration of a trial, and evident in both experiments (Exp. 1: $\beta = 2.70$, $SE = 0.45$, $t = 5.93$; Exp. 2: $\beta = 1.56$, $SE = 0.48$, $t = 3.23$). Figure 2b shows the relationship between pupil diameter and fixation saliency based on the grand mean of five pupil-diameter bins for Exp. 1. This figure shows that our partial-effect slopes (β), while based on complex analyses (as described above), reflect a real relationship that is clearly evident in the data.

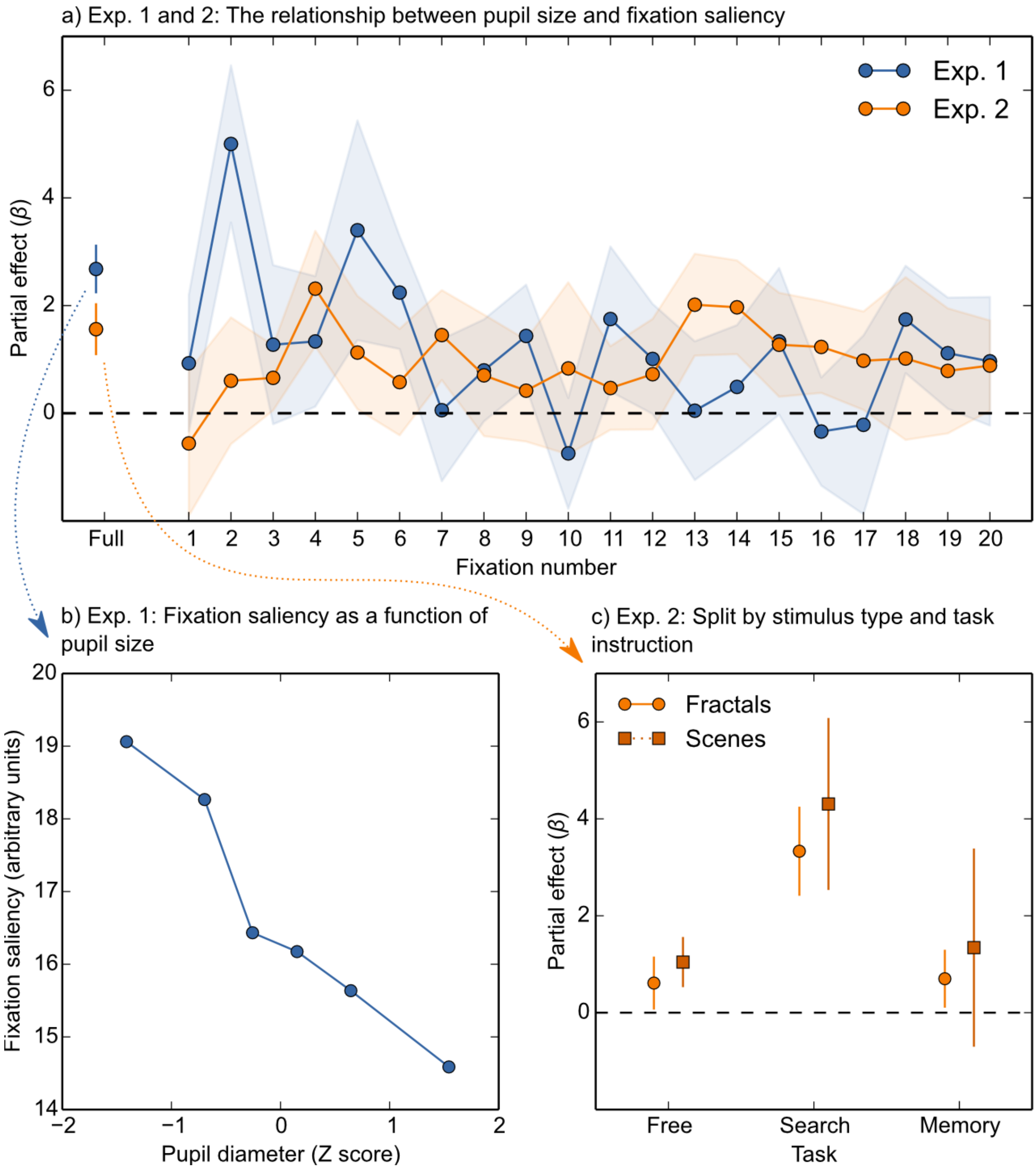

Figure 2. *The partial effect (β) of pupil size (1/D) on fixation saliency (see text for details). a) The effect for both experiments across the entire dataset ('Full') and separately for the first twenty fixations. b) The shape of the relationship between fixation saliency and pupil diameter for Exp. 1. c) The effect for Exp. 2 split by stimulus type and task instruction. Error bands/ bars indicate standard errors.*

To investigate the effect of stimulus type and task instruction in Exp. 2, we added the interaction between pupil size and stimulus type, and between pupil size and task instruction to the LME model (see Figure 2b). We also added by-participant random slopes for the effects of task instruction and stimulus type. For simplicity, we did not include a three-way interaction. This model showed that the relationship between pupil size and fixation saliency was strongest for the visual-search task (relative to free-viewing: $\beta = 2.36$, $SE = 1.01$, $t = 2.32$), with no reliable difference between the free-viewing and memory tasks ($\beta = -0.35$, $SE = 0.99$, $t = 0.36$). Furthermore, the relationship was stronger for scene than fractal stimuli ($\beta = 1.27$, $SE = 0.28$, $t = 4.57$). In other words, the relationship between pupil size and fixation saliency is generally present, but particularly pronounced during visual search and with familiar stimuli.

Temporal characteristics

To investigate whether the relationship between fixation saliency and pupil size is temporally diffuse, we repeated the analysis described above but used fixation saliency for previous and next fixations as dependent measure. In other words, we tested whether pupil size at fixation i also predicts saliency at fixation i-2, i-1, i+1, etc. This was done using the same models as in the main analyses, but varying the dependent measure (fixation-saliency-2-back, fixation-saliency-1-back, etc.).

The outcome of this analysis is striking (Figure 3): Pupil-size is strongly related to the saliency of the location that was just fixated. In general, there is a temporally diffuse window with a peak that is displaced by about one fixation. We believe that the width of this window reflects auto-correlations between measurements at different time points: Mental effort and pupil size fluctuate in cycles of several seconds (Lowenstein, Feinberg, & Loewenfeld, 1963; Reimer et al., 2014). The fact that pupil size lags behind fixation saliency may reflect the latency of the pupillary response, which can be as short as 240 ms (Beatty, 1982; Mathôt, van der Linden, Grainger, & Vitu, 2015) or as long as 700 ms (Mathôt et al., 2013), depending on a range of factors.

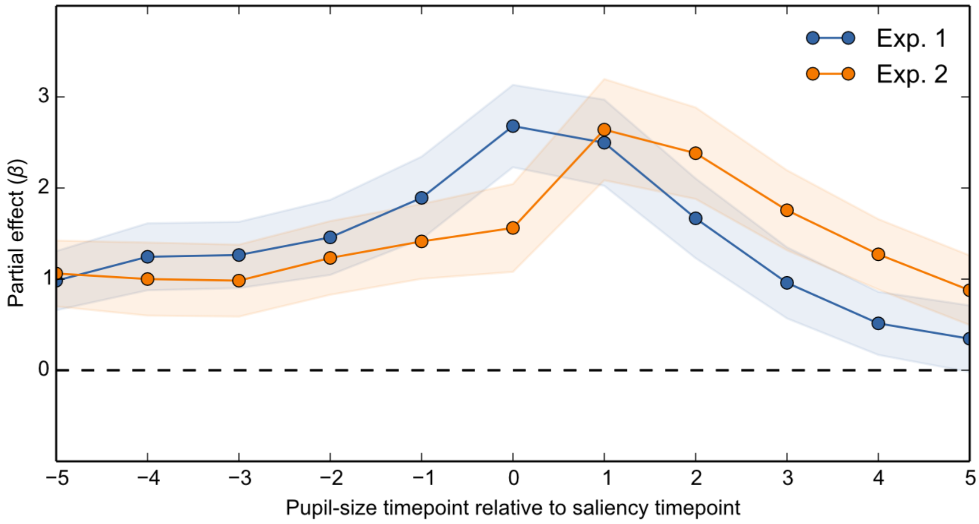

Figure 3. *The partial effect (β) of pupil size (1/D) on fixation saliency as a function of various temporal displacements. A displacement of 1 on the x-axis corresponds to the relationship between pupil size on fixation i+1 and fixation saliency for fixation i. Error bands indicate standard errors.*

Discussion

In summary, across two experiments we observed a relationship between pupil size and fixation saliency, such that you are more likely to look at low-salient locations when your pupil is relatively dilated. This relationship is not (fully) explained by any of the other variables that we considered. We interpret this relationship in terms of mental effort: Mental effort is required to engage top-down goals, and, if necessary, to overcome the inherent bottom-up bias towards high-salient locations. This interpretation is in line with the finding that the relationship between pupil size and fixation saliency was strongest in a visual-search task in which the target could be hidden anywhere, and observers therefore needed to look also at low-salient locations.

Experiment 3

The aim of Exp. 3 was to test whether the relationship between pupil size and fixation saliency is related to general arousal, or–which we think is the more interesting possibility–to mental effort that is invested specifically in guiding the eyes towards relevant locations. To this end, we compared a visual-search-only condition (cf. Exp. 1) with a dual-task condition in which the same visual-search task was secondary to an auditory memory task.

A similar dual-task logic is often used in pupilometry (e.g., Karatekin, Couperus, & Marcus, 2004). In general terms, if pupil size correlates with performance on task A, and this correlation is reduced when a second task (B) is added, then the correlation between pupil size and performance on task A is likely due to the amount of effort that is invested in task A.

More specifically, in our case, we assume that pupil size largely reflects the mental effort that is invested in the primary task. Therefore, if eye-movement control (i.e. visual search) is secondary, fluctuations in pupil size should not (only) reflect fluctuations in the effort that is invested in eye-movement control, and a large pupil should be less predictive of fixations on low-salient locations. Therefore, we predict a weaker relationship between pupil size and fixation saliency in the dual-task, compared to the single-task condition.

Methods

The methods, stimuli, and procedure were similar to those of Exp. 1, with the following exceptions. Throughout each trial, a random auditory stream of digits (1-5) was presented with an inter-digit interval of 1500 ms. In the visual-search-only condition, participants ignored the auditory stream and searched for a small ‘Z’ or ‘H’ embedded in the image (cf. Exp. 1). In the dual-task condition, the visual-search task was secondary, and the primary task was to count the number of times that a specific auditory target was presented. We were not interested in the auditory memory task per se, but it served to reduce the amount of mental effort invested in the visual-search task. In both conditions, the trial ended when the participant gave a response to the visual-search task, or after 20 s. In the dual-task condition, the auditory target was presented visually prior to the trial, and participants reported the digit count at the end of the trial. Feedback was provided only for the primary task. The set-up was similar (but not identical) to those used for Exp. 1 and 2.

Condition (dual- or single-task) switched halfway the experiment. Condition order was randomly varied between participants. The experiment consisted of 200 trials, preceded by four practice trials. Twenty-five participants were recruited.

Results

Trials lasted on average 16.1 s. The average fixation durations were 253 ms (single task) and 252 ms (dual task). Because on a small number of fixations there appeared to be recording artifacts resulting in very small pupil-size measures, we excluded fixations with a pupil size that deviated more than 3 SD from the mean pupil size (1.36 %). A total of 227,648 fixations were entered into the analysis. No participants were excluded.

First, we tested whether performance on the visual-search task was impaired in the dual-task, compared with the single-task condition (i.e. whether participants did indeed invest mental effort in the auditory memory task). To do so, we conducted an LME model with reaction time on the visual-search task as dependent variable, condition as fixed effect, and by-participant random intercept and condition slopes. This revealed a reliable effect of condition in the expected direction (β = -3.00, *SE* = 0.40, *t* = 7.58): Participants responded about 3 s faster in the single-task than in the dual-task condition.

Next, we tested whether there was an overall difference in pupil size between the two conditions. To do so, we conducted the same analysis as described above for reaction times, but using pupil size as dependent variable. This revealed that there was no notable effect of condition on overall pupil size (β = -0.03, *SE* = 0.6, *t* = 0.45), presumably indicating that participants were maximally engaged in both conditions.

For the pupil-size analysis, we used the same analysis as for Exp. 1 and 2, but added the pupil size x condition interaction as fixed effect, as well as by-participant random slopes for condition. This analysis revealed a reliable interaction in the expected direction, such that the relationship was stronger for the single-task than the dual-task condition (β = 0.43, *SE* = 0.18, *t* = 2.44). The relationship between pupil size and fixation saliency (i.e. β) was about twice as strong in the single-task (β = 0.95) than the dual-task condition (β = 0.52).

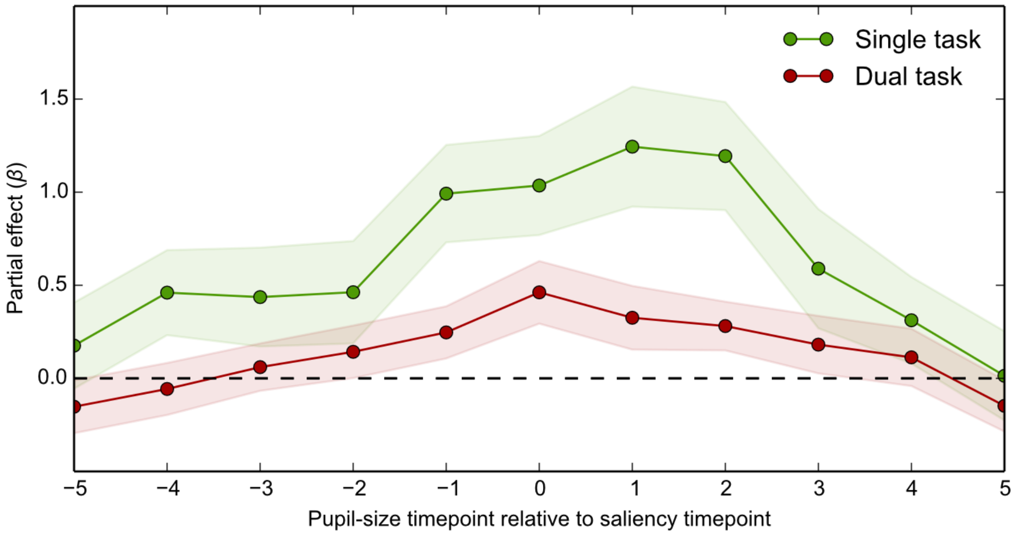

Figure 4. *The partial effect (β) of pupil size (1/D) on fixation saliency for various temporal displacements, separately for the single- and dual-task conditions from Exp. 3.*

The difference between the single- and dual-task conditions is particularly striking in the temporal-displacement analysis (Figure 4), which is identical to the analyses presented in Figure 3, but performed separately for the single- and dual-task conditions. This analysis shows clearly that in the dual-task condition the relationship between pupil size and saliency is consistently reduced for all temporal displacements.

Discussion

The single-task condition (visual-search only) in Exp. 3 replicates the negative relationship between pupil size and fixation saliency observed in Exp. 1 and 2, including the temporally diffuse pattern with a peak relationship between pupil size at fixation i+1 and saliency at fixation i. Crucially, in the dual-task condition, where the visual-search task was secondary, this relationship was much reduced.

These results support our proposal that the relationship between pupil size and fixation saliency is mediated by mental effort. When participants performed only the visual-search task, pupil size reflected the amount of effort that was invested in goal-driven eye-movement control, and large pupils therefore predicted fixations on low-salient locations. However, when the visual-search task was secondary, and pupil size partly, if not largely, reflected the effort invested in the primary task (auditory memory), large pupils became a much weaker predictor of goal-driven eye movements.

General discussion

Here we report that large pupils predict fixations of the eye on low-salient locations. We recorded unconstrained eye movements of participants who viewed images, and found a clear inverse relationship between pupil size and the bottom-up visual saliency of fixated locations. This pattern was found for different types of stimuli and under different task conditions, and held after controlling for a number of factors (luminance, saccade metrics, etc.) that might co-vary with both fixation saliency and pupil size.

We interpret this effect as showing that mental effort, reflected by a pupillary dilation (Beatty, 1982; Kahneman & Beatty, 1966; Zekveld et al., 2010), is required for goal-driven behavior. More specifically, we suggest that mental effort is required to guide gaze towards parts of a visual scene that are relevant to the task at hand, but may not be very salient. In the absence of mental effort, the eyes tend to be captured by high-salient parts of the image, regardless of top-down goals. This finding can be linked to the concept of a priority map, which is a heatmap-like representation of the visual environment that guides the visual and oculomotor systems (Bisley & Goldberg, 2010). In this map, prioritized ('hot') objects are more likely to attract the eyes. One factor that determines how strongly an object is prioritized is bottom-up visual saliency (Itti et al., 1998): An object is prioritized if it is unique (Theeuwes, 2010), has suddenly appeared (Yantis & Jonides, 1984), is moving (Franconeri & Simons, 2003), or has a biologically relevant shape (C. H. Hansen & Hansen, 1988). A person's top-down goals constitute a second factor: If you are looking for a friend in a crowded bar, people that somehow match the expected appearance of your friend are prioritized (Wolfe, 1994; Zelinsky, 2008). Crucially, our results suggest that the contribution of top-down goals to the priority map depends on the amount of mental effort that is invested, and that this contribution can be tracked through pupillometry.

Up to now, spontaneous fluctuations in pupil size have received interest mostly as an indicator of fatigue (e.g. Geacintov & Peavler, 1974). In this context, the measure of interest is usually the average amplitude of these fluctuations (increased ‘pupillary unrest’ indicates fatigue), or average pupil size (smaller pupils indicate fatigue). Only recently have researchers begun to appreciate that the different periods within a fluctuation are meaningful as well (e.g., Reimer et al., 2014). This was already foreseen by Lowenstein et al. (1963), who were among the first to systematically document fluctuations in resting-state pupil size, and conjectured that these reflected “successive waves of spontaneous arousal and following decline” (pp. 145). In line with this, and to our knowledge for the first time, the present study provides empirical support for the notion that spontaneous fluctuations in pupil size reflect fluctuations in top-down control of behavior.

For the purpose of disentangling bottom-up and top-down control, we have conveniently assumed that top-down goals necessarily drive the eyes towards low-salient locations, at least relative to a pure-saliency model of eye movements. This assumption is justified in the sense that eye movements are always guided to some extent by factors other than saliency (Tatler, Baddeley, & Gilchrist, 2005). However, this is also a simplification, because low-level saliency and top-down goals often overlap (Henderson, Brockmole, Castelhano, & Mack, 2007): Important things tend to be salient. For example, although visual saliency is not relevant to a psychology student who performs a visual-search task for course credit, it is very relevant to a rabbit that is on the lookout for predators. (Or to predators on the lookout for a rabbit.) The strongest negative relationship between pupil size and fixation saliency should be observed when bottom-up and top-down behavior can be most easily distinguished, i.e. when saliency is the least relevant to top-down goals. This is supported by the results from Exp. 2: We indeed observed the most pronounced relationship in the visual-search task, in which there is an incentive to inspect also the low-salient parts of the image, which often contained the search target. (see Figure 2c).

Our interpretation relies in part on correlations. Therefore, a legitimate question is whether different interpretations may not also account for the relationship between pupil size and fixation saliency. For example, one might conjecture that it takes additional effort to process low-salient parts of an image. This additional effort would be reflected in a dilation of the pupil, thus explaining the inverse relationship between pupil size and fixation saliency. This interpretation has the advantage that it readily explains why fixations on low-salient locations are followed, rather than preceded by a maximal dilation of the pupil (corresponding to the rightwards displacement of the peaks in Figure 2 and Figure 4), because it supposes that pupil dilation is a direct consequence of fixating low-salient locations. However, we do not favor this interpretation for two main reasons. Firstly, in the current set of images, low-salient locations typically correspond to homogeneous regions of sky, water, and grass. These are not things that one expects to require particularly intensive cognitive processing. Secondly, if low-salient locations were indeed more difficult to process, we would expect fixation durations to be longer for low-salient than for high-salient locations (Rayner, 1998), whereas we observe the opposite pattern for Experiments 1 and 3 (cf. the effect of fixation duration in Supplementary Methods, Tables 1 and 4). However, we acknowledge that different interpretations are possible. Although we have considered many confounding variables in our analyses, we cannot rule out that we have missed a crucial variable. In particular, there are factors that are difficult to quantify, such as how arousing or informative fixated locations are, that we have not considered in our analyses. It is possible that one of these factors, rather than mental effort, mediates the relationship between pupil size and fixation saliency. However, while acknowledging this limitation, we favor an interpretation in terms of top-down control of eye movements: Fluctuations in mental effort modulate pupil size as well as the tendency for the eyes to be captured by saliency. This interpretation is also supported by Experiment 3, in which we used a dual task to manipulate the amount effort in eye-movement guidance. We attribute the lag between pupil dilation and fixation saliency to the slowness of the pupillary response (Beatty, 1982).

In summary, we have shown that dilated pupils predict fixations of the eye on low-salient locations. We have interpreted this as showing that mental effort, reflected through a dilation of the pupil, is required to guide the eyes towards objects that are relevant to current goals, and to reduce capture of the eyes by high-salient, but potentially irrelevant objects. More generally, our results suggest that spontaneous fluctuations in pupil size provide an on-line marker of goal-driven behavior.

Acknowledgements

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° 622738.

References

Alnaes, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4), 1. doi:10.1167/14.4.1

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. doi:10.1016/j.jml.2007.12.005

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). doi:10.1016/j.jml.2012.11.001

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. doi:10.1037/0033-2909.91.2.276

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*, *33*(5), 2199–2204. doi:10.1523/JNEUROSCI.3440-12.2013

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, *33*, 1–21. doi:10.1146/annurev-neuro-060909-152823

Burge, J., & Geisler, W. S. (2011). Optimal defocus estimation in individual natural images.

Proceedings of the National Academy of Sciences, *108*(40), 16849–16854. doi:10.1073/pnas.1108491108

Crawford, B. H. (1936). The dependence of pupil size upon external light stimulus under static and variable conditions. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 376–395. doi:10.1098/rspb.1936.0072

Donk, M., & van Zoest, W. (2008). Effects of salience are short-lived. *Psychological Science*, *19*(7), 733–739. doi:10.1111/j.1467-9280.2008.02149.x

Foulsham, T., & Underwood, G. (2007). How does the purpose of inspection influence the potency of visual salience in scene perception? *Perception*, *36*(8), 1123. doi:10.1068/p5659

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, *65*(7), 999–1010. doi:10.3758/BF03194829

Geacintov, T., & Peavler, W. S. (1974). Pupillography in industrial fatigue assessment. *Journal of Applied Psychology*, *59*(2), 213–216. doi:10.1037/h0036529

Godijn, R., & Theeuwes, J. (2002). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1039–1053. doi:10.1037/0096-1523.28.5.1039

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *Journal of Personality and Social Psychology*, *54*(6), 917–924. doi:10.1037/0022-3514.54.6.917

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In M. Fischer, W. Murray, R. Hill, & R. P. G. van Gompel (eds.), *Eye Movements: A Window on Mind and Brain* (p. 537–562). Amsterdam, Netherlands: Elsevier.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review*, *16*(5), 850–856. doi:10.3758/PBR.16.5.850

Hong, S., Narkiewicz, J., & Kardon, R. H. (2001). Comparison of pupil perimetry and visual perimetry in normal eyes: decibel sensitivity and variability. *Investigative Ophthalmology & Visual Science*, *42*(5), 957–965.

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. doi:10.1109/34.730558

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585. doi:10.1126/science.154.3756.1583

Karatekin, C., Couperus, J. W., & Marcus, D. J. (2004). Attention allocation in the dual-task paradigm as measured through behavioral and psychophysiological responses. *Psychophysiology*, *41*(2), 175–185. doi:10.1111/j.1469-8986.2004.00147.x

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, *10*(4), 346–352. doi:10.1111/1467-9280.00166

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. doi:10.1177/1745691611427305

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–579. doi:10.1016/S0166-2236(00)01657-X

Loewenfeld, I. E. (1958). Mechanisms of reflex dilatation of the pupil. *Documenta Ophthalmologica*, *12*(1), 185–448. doi:10.1007/BF00913471

Lowenstein, O., Feinberg, R., & Loewenfeld, I. E. (1963). Pupillary movements during acute and chronic fatigue: A new test for the objective evaluation of tiredness. *Investigative Ophthalmology & Visual Science*, *2*(2), 138–157.

Marczak, K. (2012). *Mandelbulber*. Retrieved from http://mandelbulber.com/

Mathôt, S., Dalmaijer, E., Grainger, J., & Van der Stigchel, S. (2014). The pupillary light response reflects exogenous attention and inhibition of return. *Journal of Vision*, *14*(14), 7. doi:10.1167/14.14.7

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. doi:10.3758/s13428-011-0168-7

Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2013). The pupillary response to light reflects the focus of covert visual attention. *PLoS ONE*, *8*(10), e78168. doi:10.1371/journal.pone.0078168

Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2015). The pupillary light response reflects eye-movement preparation. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 28–35. doi:10.1037/a0038653

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107–123. doi:10.1016/S0042-6989(01)00250-4

Peirce, J. W. (2007). PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. doi:10.1016/j.jneumeth.2006.11.017

Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology*, *60*(2), 211–229. doi:10.1080/17470210600673818

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision*, *10*(10), e3. doi:10.1167/10.10.3

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. doi:10.1037/0033-2909.114.3.510

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422. doi:10.1037/0033-2909.124.3.372

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied*, *9*(2), 119. doi:10.1037/1076-898X.9.2.119

Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, *84*(2), 355–362. doi:10.1016/j.neuron.2014.09.033

Siebold, A., van Zoest, W., & Donk, M. (2011). Oculomotor evidence for top-down control following the initial saccade. *PLoS ONE*, *6*(9), e23552. doi:10.1371/journal.pone.0023552

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*. doi:10.1002/wcs.1323

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, *45*(5), 643–659. doi:10.1016/j.visres.2004.09.017

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*, 77–99. doi:10.1016/j.actpsy.2010.02.006

Tkačik, G., Garrigan, P., Ratliff, C., Milčinski, G., Klein, J. M., Seyfarth, L. H., … Balasubramanian, V. (2011). Natural images from the birthplace of the human eye. *PLoS ONE*, *6*(6), e20409. doi:10.1371/journal.pone.0020409

van Zoest, W., & Donk, M. (2005). The effects of salience on saccadic target selection. *Visual Cognition*, *12*(2), 353–375. doi:10.1080/13506280444000229

van Zoest, W., & Donk, M. (2006). Saccadic target selection as a function of time. *Spatial vision*, *19*(1), 61–76. doi:10.1163/156856806775009205

van Zoest, W., & Donk, M. (2008). Goal-driven modulation as a function of time in saccadic target selection. *The Quarterly Journal of Experimental Psychology*, *61*(10), 1553–1572. doi:10.1080/17470210701595555

Van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology Human Perception and Performance*, *30*(4), 746–759. doi:10.1037/0096-1523.30.4.749

Wolfe, J. M. (1994). Guided search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. doi:10.3758/BF03200774

Wolfe, J. M. (2010). Visual search. *Current Biology*, *20*(8), R346–R349. doi:10.1016/j.cub.2010.02.016

Wolff, M. J., Scholz, S., Akyurek, E. G., & Rijn, H. V. (2014). Two visual targets for the price of one? Pupil dilation shows reduced mental effort through temporal integration. *Psychonomic*

Bulletin & Review, 1–7. doi:10.3758/s13423-014-0667-5

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 601–621. doi:10.1037/0096-1523.16.1.121

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and Hearing*, *31*(4), 480–490. doi:10.1097/AUD.0b013e3181d4f251

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115*(4), 787–835. doi:10.1037/a0013118