Behavior Research Methods (2018) 50:2551–2567
https://doi.org/10.3758/s13428-018-1033-8

Using task effort and pupil size to track covert shifts of visual attention independently of a pupillary light reflex

Andreas Brocher[1,2] · Raphael Harbecke[3] · Tim Graf[1,2] · Daniel Memmert[3] · Stefanie Hüttermann[3]

Published online: 7 March 2018

Abstract
We tested the link between pupil size and the task effort involved in covert shifts of visual attention. The goal of this study was to establish pupil size as a marker of attentional shifting in the absence of luminance manipulations. In three experiments, participants evaluated two stimuli that were presented peripherally, appearing equidistant from and on opposite sides of eye fixation. The angle between eye fixation and the peripherally presented target stimuli varied from 12.5° to 42.5°. The evaluation of more distant stimuli led to poorer performance than did the evaluation of more proximal stimuli throughout our study, confirming that the former required more effort than the latter. In addition, in Experiment 1 we found that pupil size increased with increasing angle and that this effect could not be reduced to the operation of low-level visual processes in the task. In Experiment 2 the pupil dilated more strongly overall when participants evaluated the target stimuli, which required shifts of attention, than when they merely reported on the target's presence versus absence. Both conditions yielded larger pupils for more distant than for more proximal stimuli, however. In Experiment 3, we manipulated task difficulty more directly, by changing the contrast at which the target stimuli were presented. We replicated the results from Experiment 1 only with the high-contrast stimuli. With stimuli of low contrast, ceiling effects in pupil size were observed. Our data show that the link between task effort and pupil size can be used to track the degree to which an observer covertly shifts attention to or detects stimuli in peripheral vision.

Keywords Visual attention · Attentional shift · Pupillometry · Breadth of attention · Task effort

Attention, in particular visual attention, has for a long time been, and continues to be, a hot topic in cognitive psychology and related disciplines. Researchers have conducted numerous studies using many different experimental paradigms to test what attention is and does, how it is implemented in the mind, and how it is represented in the brain (Carrasco, 2011; Petersen & Posner, 2012; Posner, 1994, for reviews). In the present study, we investigated covert shifts of attention to the periphery of the visual field. Our study builds on previous work showing that the evaluation of stimuli that appear farther away from the eye's fixation is a harder task and requires more effort than the evaluation of stimuli that appear closer to the eye's fixation (Hüttermann, Memmert, & Simons, 2014; Hüttermann, Memmert, Simons, & Bock, 2013; see Hüttermann & Memmert, 2017, for a review). Because an increase in task effort has typically been associated with an increase in pupil size (see Beatty, 1982; Beatty & Lucero-Wagoner, 2000, for reviews), we reasoned that we should be able to track covert shifts of attention with pupil size when participants inspect and evaluate more distant versus more proximal stimuli. Importantly, we should be able to do so in the absence of luminance manipulations.

Understanding the link between covert shifts of attention and the size of the pupil is not only potentially interesting for researchers who aim at deciphering the attentional system; it is also relevant for researchers in the more applied sciences, such as researchers in the automotive industry and researchers in professional sports. To take an example, one characteristic of a safe car driver is that she/he can flexibly shift her/his attention to various (potentially important) locations outside foveal vision and that she/he can integrate the extracted information in her/his present estimation, or rather representation, of the ongoing traffic. Now, when developing and testing specific

✉ Andreas Brocher
abrocher@uni–koeln.de

1 CRC Prominence in Language, Luxemburger Str. 299, Room 1.06, 50939 Köln, Germany

2 Institut für deutsche Sprache und Literatur I, Universität zu Köln, Köln, Germany

3 Institut für Trainingswissenschaft und Sportinformatik, Deutsche Sporthochschule Köln, Köln, Germany

warning systems in a vehicle, covert attentional shifting becomes a relevant factor. This is because visual warning signals are unlikely to be implemented in, or close to, foveal vision. They are more likely to be implemented in a way that they appear in a driver's parafoveal vision to not disturb the driver's primary visual field. Of course, as this example also illustrates, attentional shifts are not necessarily serial. They often occur simultaneously. Covert shifts of attention therefore allow a car driver to simultaneously keep track of the speed limit, traffic signs, and pedestrians, while remaining sensitive to the sudden appearance of a visual warning signal.

That the size of the pupil can be used to measure covert shifts of attention to peripherally presented stimuli is not new. Over the past five years, there has been a body of exciting work on the relationship between the pupillary light reflex (PLR) and shifts of visual attention. One of the first demonstrations of this relationship has been presented in Binda, Pereverzeva, and Murray (2013). The authors used a variant of the Posner's cuing task in which participants were either cued to attend to a blue dot within a dark disk or a blue dot within a bright disk. Crucially, participants remained their eyes at the gray-colored center of the screen throughout the trials, with the two attended disks being equidistant to eye fixation. Participants' task was to detect subtle color changes within the blue dots that were cued. Results showed that the pupil dilated much more strongly when participants (covertly) shifted their attention to the dot in the dark disk than when they shifted their attention to the dot in the bright disk. Considering that the actual light that entered the eyes was the same in both conditions (attend to dot in dark vs. dot in bright disk), these data provide clear evidence that the PLR can be affected by higher-level cognitive processing, such as attentional shifting.

A number of studies have replicated and extended the data from Binda and colleagues (Binda, Pereverzeva, & Murray, 2013, 2014; Blom, Mathôt, Olivers, & Van der Stigchel, 2016; Mathôt, Dalmaijer, Grainger, & Van der Stigchel, 2014; Mathôt, Melmi, van der Linden, & Van der Stigchel, 2016; Mathôt, Van der Linden, Grainger, & Vitu, 2013; Naber, Alvarez, & Nakayama, 2013; Unsworth & Robison, 2017; see Mathôt & van der Stigchel, 2015, for a brief overview). Very recently, Mathôt et al. (2016) even successfully used the PLR to track participants as they selected specific letters from arrays of letters by only covertly shifting their attention to them. This successful endeavor underlines the robustness of the documented link between the PLR and covert shifts of attention.

What is important for the present study is that all of the above-cited studies exploited the PLR to track covert shifts of attention. The critical manipulation, then, always involved a dark(er) versus bright(er) distinction within the used materials. This limits the applicability of the used paradigms to stimuli that vary in luminance. Additionally, and more importantly, although the reported data clearly and unambiguously demonstrate that the pupil can reflect attentional shifts they do not address the question of whether the pupil dilates to different degrees depending on how far away the attended stimuli appear from eye fixation. Such an observation would be potentially useful for researchers who work on covert shifts of visual attention and attentional breadth, as it would allow them to not only determine *that* a person is shifting attention but also *where* in the periphery attention is being shifted to (provided that stimuli appear at different distances to eye fixation).

A study that went beyond the relationship between attentional shifts and the PLR is Gabay, Pertzov, and Henik (2011). The authors used a Posner's cuing task to investigate effects of inhibition of return on pupil size. They found that when participants responded in an identification task, judging whether they saw a "Q" or an "O" at the left or right of eye fixation, the pupil dilated more strongly than when participants responded in a localization task, judging whether or not a stimulus had appeared to the left or right of fixation. These data suggest that the evaluation of stimuli in peripheral vision leads to larger pupil dilation than the mere detection of these stimuli.

What is not clear from the data reported in Gabay et al. (2011) is what underlying process(es) led to the observed differences between the two tasks. Following considerations by James (1890) and many subsequent researchers, it is reasonable to assume that only the identification task required attentive processing. The visual appearance of the letter "Q" competed with the visual appearance of the letter "O," requiring the observer to focus on relevant and ignore irrelevant visual features of the presented stimuli. However, the identification and localization condition also differed in task difficulty and therefore in task effort: Discriminating a "Q" from an "O" is arguably more difficult and thus requires more effort than merely detecting a letter in peripheral vision. This is supported by the reported differences in response times. It took participants much longer to respond in the discrimination than the localization task. Thus, the observed differences in pupil size are likely to be a read-out of effort, with which participants mastered the task (Beatty, 1982; Beatty & Lucero-Wagoner, 2000; Kahneman, 1973; Sirois & Brisson, 2014; see also Granholm, Asarnow, Sarkin, & Dykes, 1996; Karatekin, Couperus, & Marcus, 2004; Piquado, Isaacowitz, & Wingfield, 2010; Van Gerven, Paas, Van Merrienboer, & Schmidt, 2004; Wahn, Ferris, Hairston, & König, 2016; Zekveld & Kramer, 2014). As Beatty (1982) pointed out, the "degree of pupillary dilation is a function of the processing load or 'mental effort' required to perform the cognitive task" (p. 276).

To our knowledge, there has been no systematic research on the link between pupil dilation and stimulus-to-fixation distance in the context of covert attentional shifting and task difficulty. Additionally, only few studies have investigated pupil dilation in attentional shifts in the absence of luminance manipulations. The present study was designed to close these gaps. In three experiments, we used an attention-demanding conjunction task (i.e., an attention window task; Hüttermann et al., 2013, 2014) that requires attentive processing of two stimuli presented in peripheral vision. In our experiments, the two stimuli were always equidistant to eye fixation and, crucially, appeared at various angles to eye fixation, ranging from 12.5° to 42.5°. What is important for the purpose of the present study is the observation that task performance typically decreases with increasing angle, suggesting that effort increases with increasing angle. Considering that increasing task effort is likely to also lead to increasing pupil size, pupil dilation should be stronger when participants attend and evaluate more distant stimuli than when they attend and evaluate more proximal stimuli. It is this very idea that led us to test pupil size as a proxy of covert shifts of attention in the absence of luminance manipulations.

In Experiment 1, participants counted white triangles within the peripherally presented stimuli, which also included black triangles and white and black circles. This task, then, required participants to attend to the stimuli to be able to master the task. Because counting white triangles within more distant stimuli requires more effort than counting white triangles within more proximal stimuli, we predicted larger pupils for shifts to the former than shifts to the latter. In other words, we predicted that effort in the task should increase as stimuli appear further away from eye fixation. This, in turn, should lead to stronger dilation of the pupil for more peripheral than more central stimuli.

In Experiment 2, participants either counted white triangles within the stimuli or indicated whether or not stimuli appeared in peripheral vision. This experiment was designed to test for the contribution of attention allocation to pupil size differences in the used paradigm. We expected to find overall larger pupils in the counting than the detection condition, due to differences in effort (cf. Gabay et al., 2011). The more interesting question, though, was whether the allocation of attention is necessary to yield pupil size differences between angles. It is conceivable that the mere detection of peripherally presented stimuli already yields such differences, as detecting more distant stimuli might by itself require more effort than detecting more proximal stimuli. In Experiment 3, participants always counted white triangles, like in Experiment 1. However, stimuli appeared in either high or low contrast. This contrast manipulation was included to directly test for the contribution of task difficulty and effort to pupil size differences in the attention window task, independently from angle: Counting triangles within stimuli of low contrast should be more difficult and require more effort than counting triangles within stimuli of high contrast.

Experiment 1

Method

Experiment 1 used a slightly modified version of the attention window task, in which participants evaluate briefly and peripherally presented stimuli. The attention window task has been shown to provide a robust estimate of the success with which attention has been shifted to the periphery and the effort with which the task has been mastered, typically in terms of response accuracy (see discussion below). Adjustments to the original task version (Hüttermann et al., 2013) were only made to be able to measure the size of the pupil during task performance. To that end, participants remained with their eyes at the center of a computer screen and counted white triangles that appeared within peripherally presented stimuli. Precues were briefly presented before the target stimuli to indicate to participants where on the display the target stimuli were about to appear. We tested five angles, ranging from 12.5° through 42.5°. Previous studies found that response accuracy, a proxy for task difficulty and effort, significantly decreases with increasing angle (Hüttermann & Memmert, 2015; Hüttermann et al., 2013, 2014; for a review, see Hüttermann & Memmert, 2017). The first goal of Experiment 1 was therefore to replicate these data. More importantly, however, we measured the size of the pupil as participants inspected the target stimuli. If the evaluation of more distant stimuli requires more effort than the evaluation of more proximal stimuli, we should find larger pupils when participants shift attention to the more distant than when they shift attention to the more proximal stimuli.

Participants

A total of 32 students from the University of Cologne (27 females; mean age = 22.3 years, *SD* = 3.2) participated in the experiment, either for course credit or for monetary compensation (€4). All participants had normal or corrected-to-normal vision. Two participants were excluded from the analysis because less than 60% of useable trials were recorded (Brocher & Graf, 2016, 2017; Kafkas & Montaldi, 2012, 2015). All three experiments presented here were approved by the Ethics board of the DFG as well as the Ethics board of the German Sport University in compliance with the principles of the Declaration of Helsinki 1975.

Materials

The materials consisted of precues and subsequently presented target stimuli. The precues and target stimuli appeared to the left and right of a participant's eye fixation and were equidistant from eye fixation. Participants were asked to fixate at the center of the screen throughout target evaluation. The stimuli consisted of squares that contained four objects. These objects were black or white triangles or circles, and their arrangement was randomized across trials (see Fig. 1). The size of each triangle and circle was 30 × 30 pixels. With a horizontal and vertical distance of four pixels between objects, the target stimuli had a total size of 64 × 64 pixels. The stimuli were always aligned horizontally to eye fixation, and therefore horizontally to the center of the screen. We used five different angles at which stimuli were presented from eye fixation: 12.5°, 20°, 27.5°, 35°, and 42.5°.

As precues we used small black circles that also had a size of 30 pixels. They appeared at the same locations as the subsequently presented target stimuli (see second display in sequence in Fig. 1). The precues were included so that a more low-level orienting response would not greatly interfere with any pupil dilation due to the inspection of the target stimuli.

Procedure

The experiment was conducted using a SR Research EyeLink 1000 eyetracker system (tower mounted) with a participant's head stabilized by a chin rest. The system consisted of an Intel Core i7-4770, 3.4 GHz, 4 GB RAM, running Windows 7 SP1, and a ViewSonic VS 12538 monitor. The resolution of the screen during presentation was 1,024 × 768. The distance between the eyes and the screen was about 60 cm. The experiment was programmed and presented using SR Research's Experiment Builder (version 1.6.121). The background color of all screens in the experiment was grey (RGB: 153, 153, 153) and the monitor was the only source of light in the room. The luminance measured at the right eye of participants was held at approximately 26 lux (lx) during fixation screen presentation (see below).

After they had provided written consent and after successful calibration of the eyetracker, participants were given written and oral instructions about the course of events and their task. A trial started with a fixation screen consisting of seven hash marks presented at the center of the display. The fixation screen remained for 250 ms. Subsequently, two precues (250 ms), a blank screen (200 ms), and the two target stimuli (300 ms) were presented successively (see Fig. 1). After presentation of the target stimuli, the fixation screen reappeared and was shown for 1,750 ms. This additional screen was used so that the size of the pupil could be measured in the absence of any interference from response behavior (Richer & Beatty, 1985; Richer, Silverman, & Beatty, 1983; see Brocher & Graf, 2016, 2017; Heaver & Hutton, 2011; Papesh, Goldinger, & Hout, 2012, for similar procedures). Following the fixation screen, one of two black arrows appeared. When the arrow pointed to the left, participants were asked to indicate how many white triangles they had counted within the stimulus presented to the left of fixation. When the arrow pointed to the right, participants were asked to indicate the number of white triangles they had counted within the stimulus presented to the right of fixation. Both arrows were presented at the end of a trial, and their order of appearance was randomized across trials. Participants provided their response by clicking on one of five numbers displayed on the screen that immediately followed the arrow (0, 1, 2, 3, 4), with no time pressure.

Participants started the experiment with a practice session, consisting of 15 trials. They were asked to pay close attention to the second stimuli and to count the white triangles within

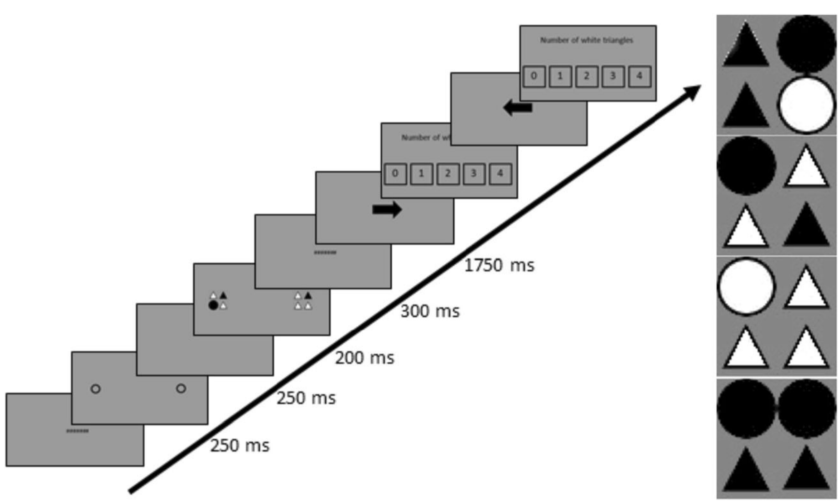

Fig. 1 (Left) Order of events in Experiment 1. At the beginning of each trial, a fixation screen was presented and quickly replaced by two precues, which indicated the locations of the subsequently presented target stimuli. A blank screen appeared between the precues and target stimuli. After the target stimulus presentation, the fixation screen reappeared, followed by a right- or left-pointing arrow (order randomized). When the arrow pointed to the right, participants reported the number of white triangles they had counted in the stimulus that appeared on the right. When the arrow pointed to the left, participants reported the number of white triangles they had counted in the stimulus that appeared on the left. (Right) Four different example stimuli for Experiment 1 (starting from top, the correct responses for the presented stimuli are 0, 2, 3, and 0).

these stimuli. Throughout practice, participants were reinstructed, if necessary, to maintain their fixation at the middle of the screen and to attend and inspect the stimuli peripherally. They were provided feedback during practice but not during the experimental session. The experimental session included a total of 150 trials, with 30 trials per angle. Each stimulus appeared equally often at each position (left or right) and angle (12.5° to 42.5°).

Recording and analysis

Recording and preprocessing Pupil size data were recorded from the right eye at 250 Hz, but vision was binocular. All blinks were excluded prior to the analysis. For each participant individually, we also excluded all pupil size data that were greater than three standard deviations from a trial's mean (Beatty & Lucero-Wagoner, 2000; Brocher & Graf, 2016, 2017; Kafkas & Montaldi, 2011, 2012, 2015). Moreover, we excluded all trials that contained fewer than 60% of useable data points (15 out of 4,051), trials in which precues or cues were overtly attended, as well as all participants that provided less than 60% of usable trials. This preprocessing procedure led to the exclusion of 5.8% of the data points and two participants.

Pupil size measure Pupil size measures were calculated as deviation from a trial's baseline and this was done for each participant and trial individually. To that end, we calculated the maximum pupil size from a trial's baseline, measured at precue presentation (250 ms), and subtracted that value from the maximum pupil size during presentation of the fixation screen following the target stimuli (1,750 ms; see Brocher & Graf, 2016, 2017; Kafkas & Montaldi, 2011, 2012, 2015, for similar procedures).

Analytic plan The data from the three experiments in this article, along with the analysis scripts, can be accessed at https://gitlab.com/tgraf0/brocher_et_al_covert_shifts_data_analysis. All statistical analyses for the three experiments presented here were conducted in R (version 3.3.1). For the behavioral data, we only classified responses as "correct" when triangles at both sides were counted correctly. Thus, the chance level for a correct response was 4% (five possible answers per side). Accuracy of participants' responses was analyzed using multilevel logistic regression models. We predicted correctness of the response by the factor angle. For pupil size measures, we fitted multilevel linear regression models. For both measures, accuracy and pupil size, we accounted for dependencies within the data by including by-participant and by-item random intercepts. For the analysis of the accuracy data, we also included a by-participant slope for angle. For the analysis of pupil size, models only converged with the intercepts, which is why the random slope was removed.

For both the analysis of responses and the analysis of the pupil size data, we fitted a null model, which only included the intercept as a predictor, and a model that included the factor Angle (12.5°, 20°, 27.5°, 35°, or 42.5°) as a predictor. We used a log-likelihood ratio test to compare the two models and determine which model fit the data better. All responses, correct and incorrect, were included in the models on pupil size deviations. Only correct responses were fed into the models of the accuracy data. The syntax of all models can be found in the Appendix.

Results and discussion

Mean response accuracy as a function of angle is presented in Fig. 2. The time courses of pupil size deviations from baseline are provided in Fig. 3. As can be seen in Fig. 2, accuracy decreased with increasing angle, replicating previous work. Not surprisingly, then, the regression model that included angle as a predictor yielded a significantly better fit to the data than did the null model, $\chi^2(4) = 55.5$, $p < .001$. More critically, and also as predicted by the effort hypothesis, pupil size increased with increasing angle, such that, again, the regression model including angle as a predictor explained the data better than did the null model, $\chi^2(4) = 72.8$, $p < .001$.

The response accuracy data were a straight replication of previous findings and showed that participants' effort in evaluating peripherally presented stimuli increased with increasing distance between these stimuli and eye fixation. The more novel finding was that the pupil also reacted to the distance at which the target stimuli were presented from eye fixation: The farther away the attended stimuli appeared from fixation, the larger the size of the pupil became. This result provides preliminary evidence that we can use pupil size to track covert attentional shifts in stimulus evaluation in the absence of luminance manipulations. We surmised that this is possible because the evaluation of stimuli farther away from the fovea (e.g., at an angle of 42.5°) is associated with more effort than the evaluation of stimuli that appear closer to the fovea (e.g., at an angle of 12.5°). Thus, when participants shift their attention to stimuli farther away from fixation in order to master the task, the pupil dilates more strongly than when they shift their attention to stimuli closer to fixation.

This assumption is fully compatible with the accuracy data, which unambiguously show that evaluating stimuli at larger angles was more difficult and therefore required more effort than evaluating stimuli at smaller angles. It is also fully in line with the data reported in Gabay et al. (2011), as well as with the many studies that have shown that increased task effort leads to increased pupil dilation (see Beatty, 1982, Beatty & Lucero-Wagoner, 2000; Kahneman, 1973; Sirois & Brisson, 2014, for reviews).

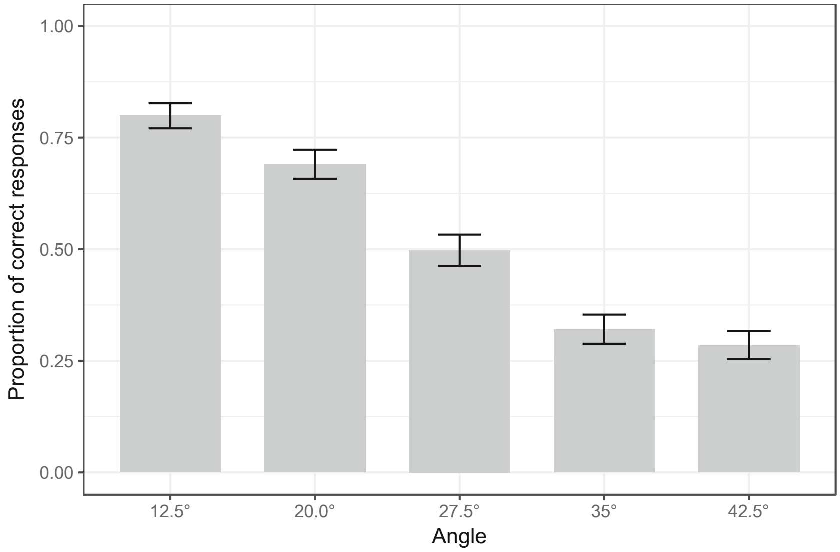

Fig. 2 Mean response accuracy per angle in Experiment 1. Bars show the proportions of correct responses. Error bars indicate the 95% confidence intervals.

Nevertheless, there is an alternative interpretation of the pupil size data of Experiment 1: The pupil dilated more strongly for more distant than for more proximal stimuli because the light associated with these stimuli hit different regions of the retina. Crucially, the number of neurons that process visual information, such as brightness, is largest at the fovea and decreases with increasing distance from the fovea. Thus, light-induced constrictions are predicted to be strongest for foveal and weakest for parafoveal illumination (Crawford, 1936). Because it is reasonable to assume that the light associated with more distant stimuli hit the retina at more peripheral regions than did the light associated with more proximal stimuli, our data might be due to the cortical magnification factor: The farther away two stimuli appear from the fovea, the fewer neurons process the light associated with these stimuli. This, in turn, would mean that, in Experiment 1, the stimuli at smaller angles led to larger light-induced constrictions, and thus weaker dilations, than the stimuli at larger angles.

Indeed, inspection of Fig. 3 supports this view. The plot reveals a constriction at around 300 ms post-stimulus-offset, and this constriction seems most pronounced for the 12.5° angle condition.

It is important to note that the actual light that entered the eyes varied very little between the numbers of white objects within the target stimuli and between angles (25.5–27.0 lx). However, we know from the literature on pupil size and the

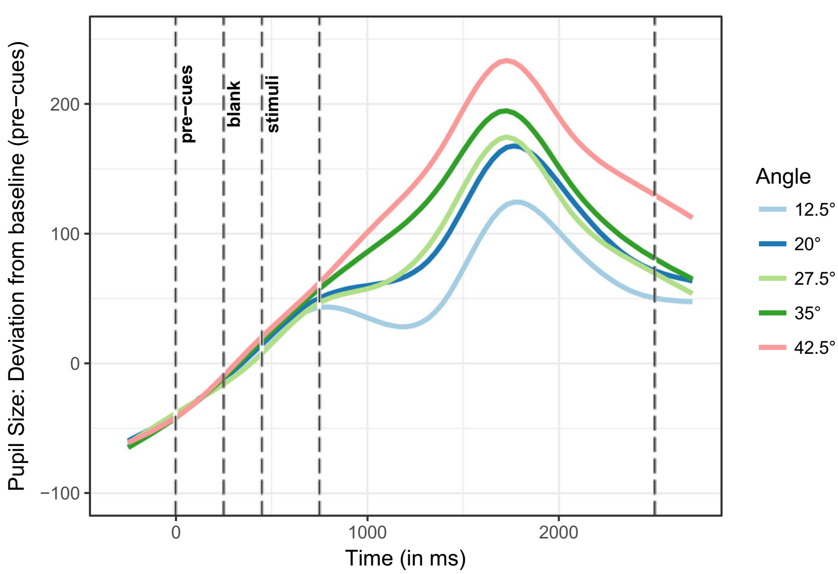

Fig. 3 Time courses for the mean pupil size deviations from baseline per angle, in number of pixels, in Experiment 1. Each angle is represented as a separate line. The dashed vertical lines represent the onsets of the precues, the blank screen, the target stimuli, and the measurement phase.

PLR that we discussed in the introduction that it is the *perceived* brightness during attentional shifts that affects the size of the pupil. We therefore addressed the potential confound between angle-related dilation and eccentricity-related constriction by conducting an additional analysis in which we included, for each trial, the number of white objects (triangles and circles) within the target stimuli. The rational is that if stimulus brightness led to stronger constrictions at smaller than at larger angles, this effect should be most pronounced with particularly bright stimuli (four white objects in the left and four white objects in the right stimulus; cf. Fig. 1), and weakest with particularly dark stimuli (four black objects in the left and four black objects in the right stimulus). In other words, eccentricity-related constrictions should be particularly strong when the light of bright stimuli hits the retina close to the fovea (e.g., with 12.5°). Such constrictions should be comparably weak when the light from bright stimuli hits the retina farther away from the fovea (e.g., at 42.5°). In contrast, the constrictions should be particularly weak, or even absent, when only little light hits the retina to begin with. Thus, for stimuli with only black objects we predict, at best, small differences between angles.

The plot shown in Fig. 4 provides the pupil size deviations from baseline for each angle as a function of the number of white objects (zero to eight). As can be seen, there is no systematic pattern between number of white objects, on the one hand, and angle, on the other. For the 12.5° and 27.5° angle conditions, there is a small decline in pupil size deviation as stimuli become brighter. For the 35° condition, this effect is somewhat stronger. However, the 20° angle condition shows virtually no effect of brightness, whereas for the 42.5° angle condition, brighter stimuli even appear to have led to larger pupil deviations than darker stimuli.

We next fitted a model that included the linear combination of angle (12.5°, 20°, 27.5°, 35°, or 42.5°) and brightness (0, 1, 2, 3, 4, 5, 6, 7, or 8 white objects) and a model that included the Angle × Brightness interaction, and compared these models to the model that included only angle as a predictor. Neither the addition of brightness as a single predictor, $\chi^2(1) = 0.27$, $p = .606$, nor the addition of the interaction term improved model fit, $\chi^2(5) = 2.92$, $p = .712$. This renders it unlikely that the data from Experiment 1 can be reduced to differences in light-induced constrictions, resulting from differences in eccentricity. The brightness of the stimuli did not systematically affect the degree to which the pupil dilated at the various angles.

After having provided evidence that the size of the pupil positively correlates with stimulus-to-fixation distance when participants shift attention to inspect peripherally presented stimuli, we turned to the question of whether attention shifting is a precondition for these effects. It is possible that the mere detection of peripherally presented stimuli is sufficient to elicit robust differences in pupil size between angles. In Experiment 2, we tested this possibility. Using the same materials as in Experiment 1, participants either counted white triangles or indicated whether or not they had noticed one or two stimuli, or no stimuli, in their parafovea. If allocation of attention to the target stimuli is necessary to elicit significant differences between angles, we should replicate the results from Experiment 1 only in the counting condition. If, on the other hand, the detection of target stimuli alone can also yield differences in pupil dilation, we expected angle to be a reliable predictor of pupil size in both the counting and the detection conditions. Finally, in line with the effort hypothesis, we predicted that the counting task would elicit overall stronger pupil dilation than the detection task, because the former should be more difficult than the latter.

Experiment 2

Method

Participants

Twenty-three participants (16 females; mean age = 23.2 years, *SD* = 2.6) took part in Experiment 2 for course credit or

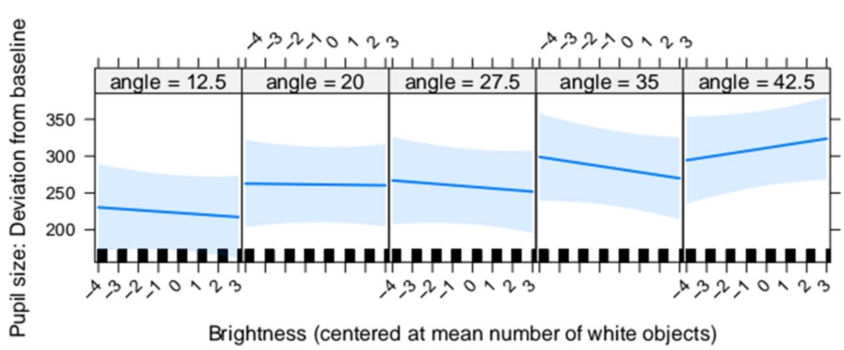

Fig. 4 Effect plot for the mean pupil size deviations from baseline as a function of the number of white objects for each angle in Experiment 1. The number of white objects appears on the *x*-axis (increasing number of white objects from left to right, centered at the mean number of white objects), and pupil size deviations from baseline appear on the *y*-axis. The different angles are represented as separate boxes.

monetary compensation (€4). All participants reported normal or corrected-to-normal vision.

Materials

We used the same precues and stimuli as in Experiment 1. However, we only included three angles: 12.5°, 27.5°, and 42.5°. Because *counting* (How many white triangles did you count?) versus *detection* (Did you notice stimuli?) was manipulated on a trial-by-trial basis, we added a prompt to the beginning of each trial, indicating to the participants which task to perform (count or detect). The prompt in the counting condition displayed the word "Count," and the prompt in the detection condition displayed the words "Yes/No."

Procedure

At the beginning of a trial, participants pressed a button to signal that they understood what task to perform (i.e., count or detect). After that and up until the presentation of the first response arrow, the course of events was the same as in Experiment 1, with the only exception being that the first fixation screen lasted for 500 ms instead of 250 ms (cf. Fig. 1). In the counting condition, participants again clicked on the number on the screen that represented their response (zero to four triangles). In the detection condition, they pressed the left button on the mouse when they thought that no stimulus had appeared on the side the arrow pointed to ("no" response), and the right button on the mouse to indicate that a stimulus had appeared on that side ("yes" response). As in Experiment 1, the arrows appeared successively, and the order of direction (left then right or right then left) was randomized across trials.

Before the main part of the experiment, a practice session of ten trials was provided. The experimental session included a total of 90 trials, with 15 trials per angle in the counting and an additional 15 trials per angle in the detection condition. Number of triangles was counterbalanced across sides (left and right) and angles (12.5°, 27.5°, and 42.5°) in the counting condition. In the detection condition and for each angle, four trials included stimuli on both sides, four trials a stimulus only on the left side, four trials a stimulus only on the right side, and three trials no target stimulus.

Recording and analysis

Recording and preprocessing We followed the same recoding and preprocessing procedure as for Experiment 1. One trial was excluded (out of 2,070 trials). A total of 4.1% of the data did not enter the statistical analysis. Note that the chance level of responding correctly to both sides in a trial was 4% (five possible answers per side = 25 possibilities in total) in the counting and 25% (present vs. absent per side = four possibilities in total) in the detection condition.

Pupil size measure The calculation of pupil size deviations was the same as for Experiment 1.

Analytic plan Accuracy scores were again analyzed using multilevel logistic regression. We only considered those responses as "correct" for which both sides were responded to correctly. We predicted the correctness of responses by the factors Angle (12.5°, 27.5°, or 42.5°) and Task (counting or detection). For peak pupil size, we used multilevel linear regression and tested the effects of angle and task on peak pupil size. The model-fitting procedure was similar to the procedure reported for Experiment 1. However, due to the addition of a second predictor, we now compared a total of four models: an intercept-only model, a model that only included angle as a predictor, a model that included the linear combination of angle and task as a predictor, and a model that additionally included the interaction of angle and task as a predictor. We again used log-likelihood ratio tests to determine which model best explained the data. The syntax of all models can be found in the Appendix.

Results and discussion

The accuracy data are shown in Fig. 5, and the time courses of peak pupil size are plotted in Fig. 6. Note first that the accuracy data in the counting condition nicely replicate the data from Experiment 1: Accuracy decreased with increasing angle. For the detection condition, performance was at ceiling, which is not surprising, considering that detecting briefly presented stimuli in the periphery is an easy task. As for statistical significance, the model that included the interaction term of angle and task yielded the best fit, $\chi^2(2) = 9.6$, $p = .008$.

Turning to the pupil size data, we again found that the counting condition replicated the data from Experiment 1: Pupil size deviations increased with increasing angle. We then note the overall steeper increases in pupil size deviation in the counting than in the detection condition. This aspect of the data was also expected, considering that discriminating white triangles from black triangles as well as from white and black circles is a harder task than judging whether or not stimuli had appeared on the screen. This pattern also nicely aligns with the finding in Gabay et al. (2011) that discriminating a "Q" from an "O" yields larger pupils than does merely providing the location of a letter (left or right). With respect to the main question of Experiment 2, Fig. 6 suggests that both the detection and counting conditions led to larger pupils for more distant than for more proximal stimuli, although this effect seems somewhat more robust for the counting condition.

Statistical analyses confirmed that pupil size increased with increasing stimuli-to-fixation distances. The model that included angle as a predictor was a better fit to the data than the intercept-only model, $\chi^2(2) = 10.4$, $p = .005$. The additional inclusion of task further improved the model fit, $\chi^2(1) =$

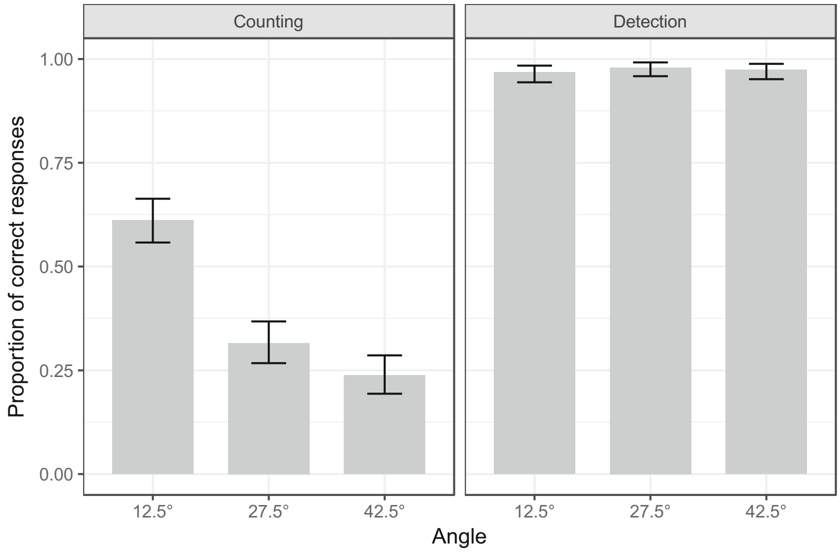

Fig. 5 Mean response accuracy per angle and 95% confidence intervals for the counting (left) and detection (right) conditions in Experiment 2.

23.9, $p < .001$, which statistically confirms that pupil size also increased more strongly overall in the counting than in the detection condition. The model that included the Angle × Task interaction term did not yield a better fit to the data than the model that included these factors as linear predictors, $\chi^2(2) = 2.9$, $p = .231$.

The results of Experiment 2 suggest that shifting attention to peripherally presented stimuli—for example, to evaluate specific properties of these stimuli—might not be a necessary precondition for increased pupil dilation as a function of stimuli-to-fixation distance. Just as in Experiment 1, the pupil dilated more strongly for more distant than for more proximal stimuli, irrespective of whether participants counted white triangles or reported whether or not any stimuli had appeared onscreen. This finding is potentially important for research on vision, because it demonstrates that pupil size can be used to test for both the detection of and attentional shifts to stimuli in peripheral vision. The observation that the counting condition, which required attentional shifts, led to larger pupils overall than did the detection condition additionally suggests that pupil size may distinguish a scenario in which the participant detects stimuli from a scenario in which the participant attends to the stimuli. Indeed, visual inspection of Fig. 6 reveals that the differences between the three angles are more robust in the counting than in the detection condition. In fact, the time courses for the different angles of the detection condition

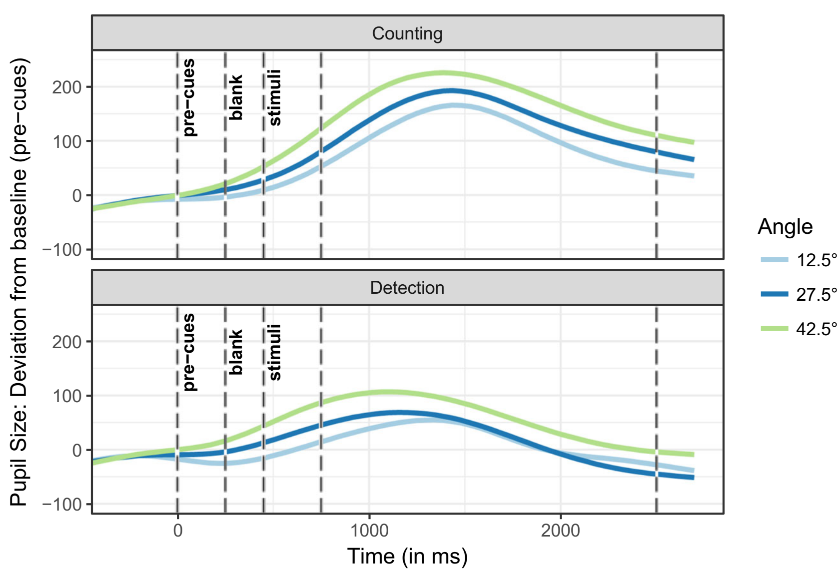

Fig. 6 Time courses of the mean pupil size deviations from baseline per angle and task, in number of pixels, in Experiment 2. Each angle is represented as a separate line. The dashed vertical lines represent the onsets of the precues, the blank screen, the target stimuli, and the measurement phase. The counting condition is plotted in the top panel, and the detection condition at the bottom.

show strong differences in pupil size between the 42.5° angle, on the one hand, and the 12.5° and 27.5° angles, on the other, whereas the two smaller angles did not greatly differ from one another.

The results of the detection condition are a good illustration of what Kahneman (1973) referred to as "effectiveness" and "efficiency." Kahneman proposed that effectiveness is mirrored in the quality of behavioral performance, whereas efficiency links this performance to the amount of resources invested in the task: The more resources a participant needs to invest to reach some level of performance, the less efficient she or he is. Now, in the detection condition, we observed almost perfect performance in the behavioral data, which means that our participants were highly effective. However, the pupil size data reveal that efficiency decreased with increasing angle. It seems that participants needed to invest the most resources to detect stimuli in the 42.5° angle condition, and the least resources in the 12.5° condition (see also Karatekin et al., 2004, for comparable results from dual-task performance). The detection condition therefore shows that pupil size may be useful for uncovering specific processes that underlie performance but that cannot easily be isolated via response behavior.

So far we have followed Kahneman (1973), Beatty (1982), and others in their assumption that there is a tight link between attention and task effort, on the one hand, and task effort and pupil size, on the other. We therefore proposed that the pupil size differences reported for Experiments 1 and 2 were due to differences in task effort: Evaluating (and detecting) stimuli farther away from the fovea was a harder task than evaluating stimuli closer to the fovea. Because the evaluation of peripherally presented stimuli goes hand in hand with covert shifts of attention, since it requires covert shifts of attention, we were able to make use of pupil size to track these shifts.

Our assumption is fully in line with the negative correlation of response accuracy and pupil size that we found across the two experiments, including the overall larger pupils in the counting than in the detection condition of Experiment 2. However, one shortcoming of Experiments 1 and 2 is that they do not provide direct and unambiguous evidence for the effort hypothesis, because task effort and shifts of attention were necessarily confounded. In both experiments, evaluating more proximal stimuli (easier task) versus more distant stimuli (harder task) was confounded with smaller versus larger shifts of attention. Furthermore, in Experiment 2 the presumably easier task—that is, the detection task—did not require attentional shifts, whereas the presumably harder task—that is, the counting task—did.

In Experiment 3, we set out to provide independent evidence for the claim that our task involved different degrees of effort and that this very observation was what allowed us to track covert shifts of attention involved in the task. This experiment was identical to Experiment 1, with the exceptions that we used only three angles and that the target stimuli appeared in either high or low contrast. Presenting stimuli in low rather than high contrast should make it harder for participants to discriminate between triangles and circles, as well as between the "white" and "black" objects. This was expected because objects in low contrast appear to have fuzzier edges and, additionally, are overall more difficult to separate from the background color of the screen (see Fig. 7). Critically, the manipulation of contrast allowed us to investigate differences in task effort independently from differences in attentional shifts.

Note that the direct manipulation of effort is also important to establishing potential boundary conditions for the use of pupil size in research on covert shifts of attention. It has been shown that some maximum amount of effort can lead to ceiling effects in the size of the pupil (Wahn et al., 2016), or even to an abrupt drop-off of pupil size, indicating task disengagement (Granholm, et al., 1996; Van Gerven et al., 2004; Zekveld & Kramer, 2014).

For Experiment 3, and in line with our effort hypothesis, we predicted larger dilations in the low- than in the high-contrast conditions, in addition to overall larger pupils for more distant than for more proximal stimuli.

Experiment 3

Method

Participants

Twenty-four participants took part in Experiment 3 (18 females; mean age = 25.1 years, *SD* = 7.7), either for course credit or for monetary compensation (€4). All participants had normal or corrected-to-normal vision.

Materials

We used the same materials as in Experiment 1, with two exceptions. First, we added a low-contrast condition. In the high-contrast condition, the target stimuli were presented at full contrast, as had been the case for Experiment 1 (RGB for white objects: 254, 254, 254; RGB for black objects: 0, 0, 0). In the low-contrast condition, the target stimuli were presented with RGB (200, 200, 200) for "white" objects and RGB (50, 50, 50) for "black" objects. An example of the stimuli used is provided in Fig. 7. Second, we used only the three angles 12.5°, 27.5°, and 42.5°.

Procedure

The procedure was identical to that of Experiment 1, with the exception that the first fixation screen lasted 500 ms instead of 250 ms (cf. Fig. 1).

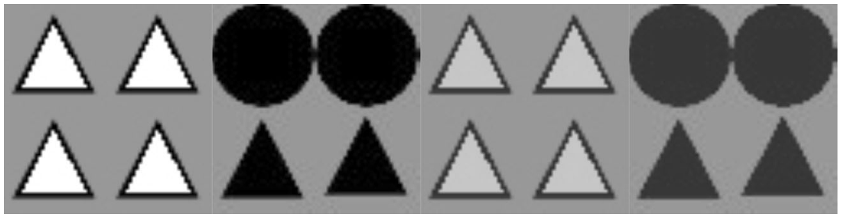

Fig. 7 Example stimuli used in Experiment 3. The two high-contrast stimuli appear on the left, and the two low-contrast stimuli appear on the right.

Recording and analysis

Recording and preprocessing The recording and preprocessing were the same as for Experiments 1 and 2. We excluded one trial (out of 2,160). A total of 4.1% of the data did not enter the analysis.

Pupil size measure The calculation of peak pupil size was the same as for Experiments 1 and 2.

Analytic plan We followed the same model-fitting and model comparison procedures as for Experiments 1 and 2. The two predictors were angle (12.5°, 27.5°, or 42.5°) and contrast (low or high). All models are provided in the Appendix.

Results and discussion

The accuracy data for Experiment 3 are displayed in Fig. 8. The time courses of pupil size deviations are plotted in Fig. 9. As in Experiment 1 and the counting condition of Experiment 2, we found decreasing accuracy with increasing angle, which was true for both the low- and high-contrast conditions. However, Fig. 8 also reveals that accuracy was lower overall in the low- than in the high-contrast condition. This confirms that the contrast manipulation was successful, in that counting "white" triangles among the low-contrast target stimuli was a more difficult task and required more effort than counting the actual white triangles amid the high-contrast target stimuli. Not surprisingly, the model that included the linear combination of angle and contrast yielded the best fit to the data, $\chi^2(2) = 53.1, p < .001$. Inclusion of the Angle × Contrast interaction did not further improve the model fit, $\chi^2(2) = 0.7, p = .717$. In sum, we observed two independent sources of difficulty: stimuli-to-fixation distance and stimulus contrast.

Turning to the pupil size data, inspection of Fig. 9 reveals that the high-contrast condition replicated the results of Experiment 1 as well as the results of the counting condition of Experiment 2. This was expected, since these conditions used the very same materials and design. We then note that the low-contrast condition yielded ceiling effects in pupil size deviation. In other words, in the low-contrast condition, the three angles led to similarly large pupil dilations, which, in turn, were comparable to the dilations with the 42.5° angle in the high-contrast condition. Indeed, the regression model that included the Angle × Contrast interaction explained the data better than the model that included angle and contrast only as linear predictors, $\chi^2(2) = 6.1, p = .048$.

The results of Experiment 3 provide direct support for the claim that changes in pupil size in a task involving covert attentional shifts reflect differences in the effort with which that task is mastered. In particular, in light of the tight link between response accuracy, on the one hand, and pupil size, on the other, we propose that effort is a strong predictor of pupil dilation in tasks that require participants to covertly shift their attention. Of course, this does not rule out the possibility that other processes might affect pupil size when attention is shifted as well. For example, it is conceivable that attentional shifts per se can involve different degrees of effort, with larger shifts requiring more effort than smaller shifts (see the General Discussion). At a minimum, however, our data demonstrate that pupil size in covert shifts of attention is sensitive to task difficulty and effort.

An important remark is in order before we continue. One might argue that the low-contrast condition in Experiment 3 yielded a ceiling in pupil size because the materials in this condition were too dark to allow for any additional dilation triggered by the task. That is, although the actual differences in brightness between the angles and the stimuli were small (25.5–27.0 lx), the observed ceiling effects might instead have been due to perceived differences in brightness. Indeed, at least for the "white" objects among the stimuli, this might in fact have been the case (cf. Fig. 7). However, Fig. 7 also shows that the stimuli with four "white" objects in the low-contrast condition should appear as bright as, or even brighter than, the stimuli with four black stimuli in the high-contrast condition. Inspecting the former should then have led to the same amount of, or even to less, dilation than inspecting the latter, assuming that differences in brightness led to the differences in pupil size.

We therefore compared particularly bright stimuli in the low-contrast condition with particularly dark stimuli in the high-contrast condition for all three angles. In Fig. 10, we plot the mean pupil size deviations from baseline for all stimuli with zero or one "white" object and for all stimuli with seven or eight "white" objects.[1] The plot reveals slightly larger pupil

[1] Extending the analysis beyond the extreme cases—that is, the stimuli with only zero or eight "white" objects—allowed us to include a sufficiently large number of data points (total of 1,032).

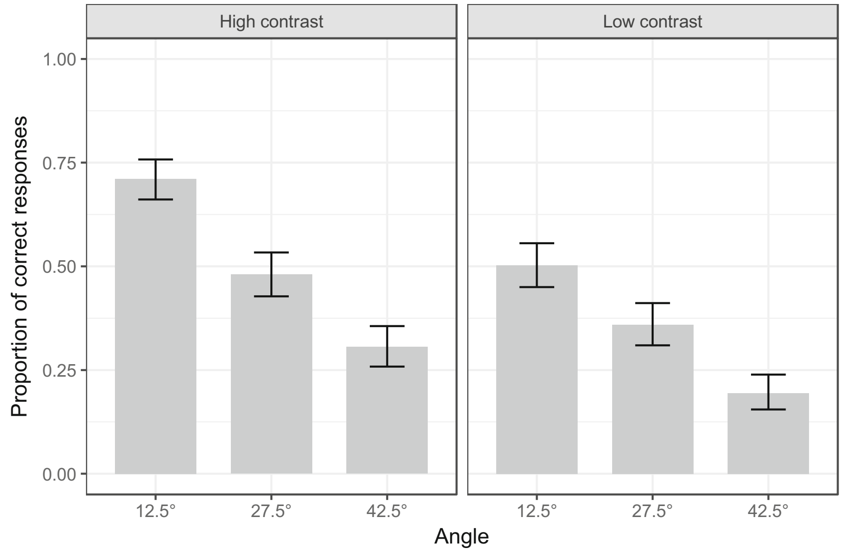

Fig. 8 Mean accuracy scores, along with 95% confidence intervals, for the high-contrast (left) and low-contrast (right) conditions in Experiment 3.

deviations for high-contrast stimuli with predominantly black objects than for low-contrast stimuli with predominantly "white" objects only for the 42.5° angle. For the 12.5° angle, at which we would expect the largest effect of brightness (see the discussion in Exp. 1), the plot even reveals the opposite pattern.

Note that, for the 12.5° angle condition, we found a pattern of results similar to that in the 12.5° angle condition of Experiment 1 (cf. Fig. 4): Stimuli with many "white" objects led to slightly smaller pupils than did stimuli with many "black" objects, and this observation held equally for both the high- and low-contrast conditions of Experiment 3. This shows that the pupil was in fact somewhat sensitive to differences in brightness, at least when the distance between eye fixation and the stimuli was small. When all conditions are taken together, though, the data from the low-contrast condition are difficult to explain under a brightness explanation. It seems difficult to explain why the particularly bright stimuli in the low-contrast condition failed to yield overall smaller pupil sizes than the particularly dark stimuli in the high-contrast condition. We argue, then, that the means in Fig. 10 are incompatible with the view that the differences in Experiment 3 can be reduced to differences in brightness.

Instead, we propose that the ceiling in pupil size for the low-contrast condition was due to exceeding a maximum threshold of effort (cf. Wahn et al., 2016). Put differently, the

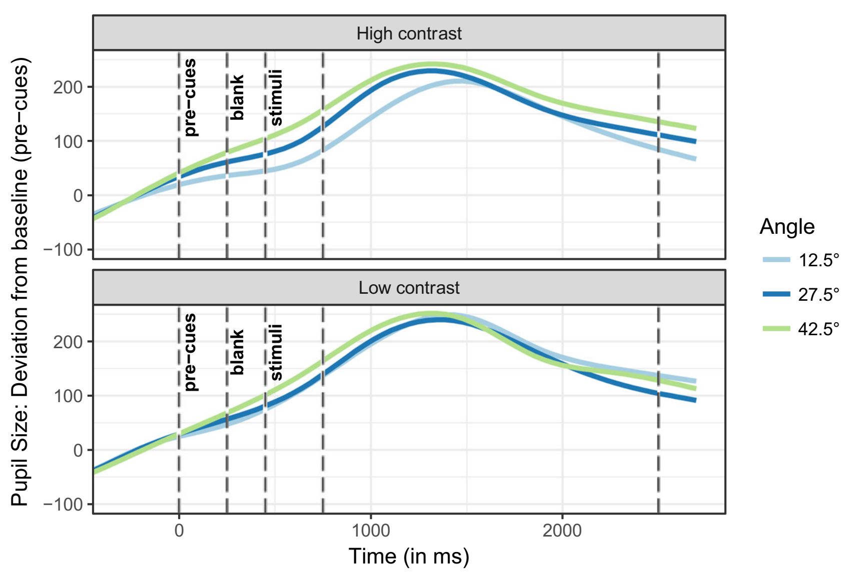

Fig. 9 Time courses of the mean pupil size deviations from baseline per angle and condition, in number of pixels, in Experiment 3. Each angle is represented as a separate line. The dashed vertical lines represent the onsets of the precues, the blank screen, the target stimuli, and the measurement phase. The high-contrast condition is plotted in the top panel, and the low-contrast condition at the bottom.

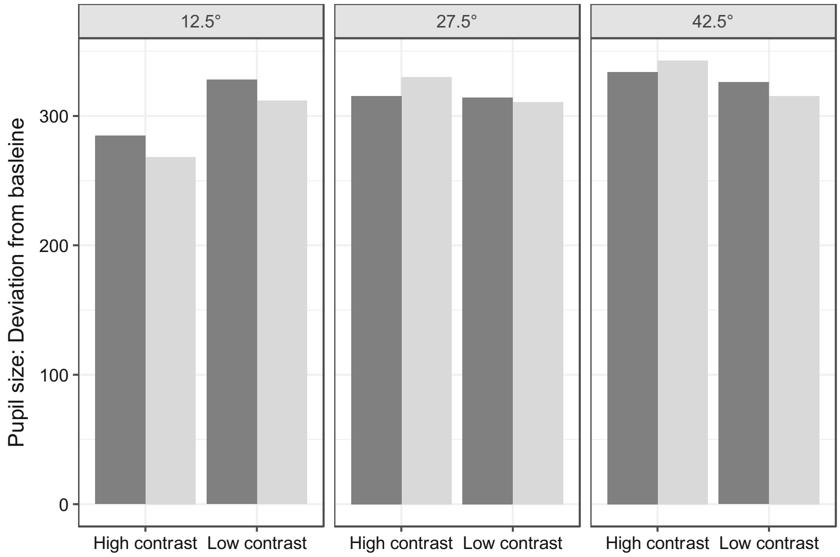

Fig. 10 Mean pupil size deviations from baseline for dark (zero or one "white" object) and bright (seven or eight "white" objects) target stimuli, per angle and condition in Experiment 3. The dark stimuli are represented by dark gray bars, and the bright stimuli are represented by light gray bars.

pupil does not continue to dilate once a participant's effort exceeds some maximum, provided that the participant has remained engaged in the task. With the latter point in mind, it is important to remember that the participants in Experiment 3 were in fact overall engaged in the task. They performed much better than chance throughout, including the presumably most difficult condition—that is, the 42.5° condition with low-contrast target stimuli (~20% correct).

General discussion

In the present study we asked two questions: Can we use the size of the pupil to track covert shifts of attention independently of a pupillary light reflex, and if so, does pupil size change as a function of the distance between eye fixation and attended stimuli? In three experiments, we used a version of the attention window task, which (a) has proven a powerful tool to assess a person's breadth of attention, because (b) the evaluation of stimuli farther away from eye fixation requires more effort than the evaluation of stimuli closer to eye fixation (see, e.g., Hüttermann & Memmert, 2017, for a review). Considering the well-documented positive correlation between task effort and pupil size, the attention window task then lent itself to testing different degrees of covert shifts of attention that are involved in the evaluation of the target stimuli by using pupil size as a metric.

In our study, participants counted white triangles within two simultaneously presented stimuli that also contained black triangles as well as black and white circles. This task, therefore, required participants to focus on relevant visual features of the stimuli while ignoring or suppressing irrelevant features (i.e., shape and/or color). Thus, the task we used required participants to allocate attention to the presented materials. Importantly, the target stimuli were presented on opposite sides of central fixation, with angles ranging from 12.5° to 42.5°. The luminance of the materials used was held constant between the different angles by using the very same materials across angles.

In all three experiments we found that the pupil dilated when participants covertly shifted their attention to the target stimuli, and that these dilations increased with increasing distance from the attended stimuli to eye fixation. Post-hoc analyses confirmed that, although at small angles there was a slight tendency for brighter stimuli to yield larger constrictions than darker stimuli, the observed differences in dilation could not be reduced to these (small) differences in constrictions. In other words, the critical differences in dilation are unlikely to have resulted from eccentricity-induced constrictions (Exp. 1) or differences in stimulus brightness (Exp. 3).

In Experiment 2, participants either counted white triangles in the target stimuli or reported whether or not these stimuli were present. Our results show that counting white triangles, a task that requires the allocation of attention, led to larger pupils overall than detecting the target stimuli, a task that presumably does not strongly engage the attentional system. Interestingly, however, even in the detection task, the pupil dilated more strongly for more distant than more proximal stimuli. This suggests that the allocation of attention to the target stimuli might not be necessary to yield larger pupils for more distant than more proximal stimuli. Nevertheless, visual inspection of the respective time courses of pupil size deviations called for some caution. For the counting condition, pupil size successively increased with increasing angle. For the detection condition, in contrast, there was a split between the small and medium angles (12.5°, 27.5°), on the one hand, and the large angle (42.5°), on the other. Although this difference between the counting and detection conditions was not statistically reliable, the descriptive data suggest that pupil size differences between the three angles were generally more robust in the counting than in the detection condition.

More conclusive seems the observation that the pupil dilated much more strongly overall in the counting than in the detection task, presumably because of differences in task difficulty. This suggests that the size of the pupil discriminates between a condition in which participants need to attend to the target stimuli from a condition in which the allocation of attention to the stimuli is not a prerequisite for mastering the task. Nevertheless, clearly, additional experiments are needed to further test for potential differences and similarities between pupil dilation in covert shifts of attention and pupil dilation in stimulus detection.

One question that arises from our data is what underlying process(es) led to the reported differences in pupil dilation. Considering the large body of literature that discusses the close relationship between task effort and pupil size (e.g., Gabay et al., 2011; Granholm et al., 1996; Granholm & Verney, 2004; Karatekin et al., 2004; Piquado et al., 2010; Van Gerven et al., 2004; Wahn et al., 2016; Zekveld & Kramer, 2014; see Beatty, 1982, for a discussion), we surmise that our data were most likely due to differences in task difficulty and effort. The farther away two stimuli appeared from eye fixation the more difficult it was for participants to evaluate or detect these stimuli, resulting in increased pupil dilation. The accuracy data from the three experiments fully corroborate this claim. We found a general increase in pupil size with decreasing accuracy. Arguably, response accuracy is strongly related to task difficulty and effort.

However, our results from Experiment 2 suggest that this relationship is not perfect. Although, again, poorer performance generally mapped onto larger pupils, this was not true for the condition, in which participants reported on the presence versus absence of peripherally presented stimuli. Here, we observed almost perfect performance throughout angles, whereas, as we stated above, pupil size increased more for the 42.5° angle than for the two smaller angles (12.5°, 27.5°). We argued that this finding supports Kahneman's (1973) distinction of effectiveness versus efficiency: Although participants were highly effective in managing their task, the size of the pupil revealed that participants were less efficient in reaching their high performance in the 42.5° than in the 12.5° and 27.5° angle conditions. In other words, reaching near ceiling performance in the large angle condition was more effortful and required more resources than reaching near ceiling performance in the smaller angle conditions.

We then tested the effort hypothesis more directly in Experiment 3, in which the stimuli were presented in either high or low contrast, both within and across angles. The rationale was that the presentation of target stimuli in low contrast should make the discrimination of the different objects within these stimuli more difficult, which was indeed what we found in the accuracy data. More interestingly, though, whereas the high-contrast condition fully replicated the counting data from Experiments 1 and 2, the low-contrast condition led to ceiling effects in pupil size. These ceiling effects, we propose, were due to the exceeding of a maximum threshold of effort. That is, the pupil seems to only reflect task effort up to a maximum amount. When that maximum is reached, no further dilation occurs. Support for this hypothesis comes from data that were recently presented by Wahn et al. (2016). The authors measured the size of the pupil as participants engaged in a multiple object tracking task. Their results show a steep increase in pupil size when the number of to-be-tracked objects increased from one to two as well as when it further increased from two to three. However, this pupil size increase flattened considerably when a fourth and fifth object were added to the set of to-be-tracked objects. We suggest that our data from Experiment 3 reflect the very same processes. Task difficulty and effort increased successively with increasing distance from target stimuli to eye fixation and, when the discrimination between white triangles and competing objects became particularly difficult, no additional dilation occurred.

This finding from Experiment 3 is important because it shows that pupil size can only be used to track covert attention shifts as long as the task difficulty does not exceed a maximum threshold of effort. When the task becomes too difficult, shifts of attention can no longer be assessed. An interesting question for future research is what properties of attentional shifting or of the test materials can lead to ceiling effects in pupil size.

Before we continue with potential applications of the present findings in specific fields of research, we would like to point out that our data do not allow us to determine whether shifts of attention themselves—that is, independently of task difficulty and task effort—lead to larger versus smaller pupil dilation. That is, it is conceivable that the mere shifting of attention to stimuli in peripheral vision is more effortful, and thus yields larger pupils, when the distance between the fovea and the to-be-evaluated stimuli is larger than when it is small. Although in particular our data from Experiment 3 provide strong evidence that task effort *can* affect pupil size independently of the size of attentional shifts, our study provides no data for a context in which shifts could potentially lead to differences in pupil size independently of task effort.

One way to test this latter scenario would be to hold task difficulty and effort constant across stimuli-to-fixation distances so that participants would be similarly successful in their response behavior across angles. If shifts of attention per se require some effort and if larger shifts require more effort than smaller shifts, we should find larger pupils when attention is shifted to stimuli father away from fixation than when it is shifted to stimuli closer to fixation, despite the fact that both angles would involve the same amount of task effort.[2]

Taken together, the results of the present study are useful for both more theoretical and more applied approaches to attentional shifting. In the remainder of this article, we discuss potential extensions of our findings. For all these extensions, it is important to keep in mind that the implemented task should involve some minimum amount of effort as well as different degrees of effort. Although this might seem somewhat obvious, as any task that requires covert shifts of attention is likely to also involve some measurable amount of effort, one should make sure that different degrees of shifts (e.g., large vs. small) line up with different amounts of task effort, just like we did in our experiments. For example, without further research it is not clear whether we could measure different degrees of covert shifts of attention if we were to hold the amount of effort between the various shifts constant.

Turning to the first potential application of our findings, we used a paradigm that was conceptually very similar to the attention window task developed by Hüttermann and colleagues (Hüttermann & Memmert, 2015; Hüttermann et al., 2013, 2014, see also Hüttermann & Memmert, 2017, for a review). However, in all previous studies that used the attention window task, the assessment of participants' breadth of attention was entirely based on the accuracy with which participants performed the task. High response accuracy signaled predominantly successful shifting, whereas low accuracy was taken to indicate failure in shifting. This means that experimenters of previous studies needed to explicitly direct a participant's attention to the peripherally presented stimuli. The data presented here open the possibility that a person's breadth of attention can be assessed without the need to explicitly instruct participants where and what to attend to. Put differently, our results suggest that a participant's breadth of attention can be measured in conditions in which the participant can freely choose when to shift attention and where to shift it to. In the counting conditions of all three experiments, we found a strong negative correlation between response accuracy and pupil size with high-contrast materials. This means that the recording of pupil size deviations seems sufficient to measure (degrees of) covert shifts of attention.

One way to implement this idea could be to simultaneously present multiple stimuli onscreen and to separate the experiment into forced-choice and free-choice trials. In forced-choice trials, participants would be asked to attend stimuli at various locations on the display and report some property of these stimuli. In the critical free-choice trials, however, participants would be free to choose which stimulus or stimuli on the screen to attend to. By comparing the pupil size deviations in the forced-choice and free-choice trials, researchers could infer which stimuli participants were attending to when their behavior was not directed. Such a paradigm—which, of course, needs to be further developed and tested—would allow researchers to manipulate specific properties of the task as well as the test materials and to assess the extent to which covert shifts of attention are affected by these manipulations.

Our data will also be useful for more applied approaches to covert shifts of attention. Returning to the traffic example from the introduction, the present findings suggest that covert shifts of attention could be measured while a person is engaged in a specific task, such as when maneuvering a vehicle (in a driving simulator). This seems very useful when developing and testing specific systems that assist a driver or warn her/him when a potentially dangerous situation is developing. As we discussed above, visual signals of assistance are likely to be implemented in a driver's parafoveal vision, such that it does not interfere with her/his foveal vision. Nevertheless, these signals need to be implemented in such a way that the driver can covertly shift attention to them and evaluate their meaning, so that the current driving behavior can be updated accordingly. For example, assistance during driving or parking might involve the display of different colors or shapes, and changes in the signal could indicate to the driver how close the car is to other cars or objects in the immediate environment. Likewise, warning signals that appear in the visual periphery could help with evaluating blind spots, indicating whether—and if so, where in a driver's blind spot—a car is approaching.

A final potential implementation pertains to professional sports. Assessing whether or not a player in a game is shifting attention at a specific time would be useful for research on visual attention in team sports. For example, a soccer player needs to be able to continuously assess to whom in the field she or he can pass the ball. Arguably, this evaluation often occurs in parafoveal vision, since the player also needs to track the ball as well as the opponent players close to her or him. Now, successful assessment of the current passing options will involve not only knowledge about where in the field the team mates are, but also an accurate assessment of the players' speeds and whether or not they are covered by opponent players. In other words, successful team sports critically involve successful shifts of attention. The link between covert shifts of attention and pupil size might help researchers develop strategies that improve a player's shifting behavior and determine which properties of the game make shifting easier or harder.

Conclusions

In the present study, we found evidence that the effort involved in a task that requires covert shifts of attention to peripherally presented stimuli leads to pupil dilation, and that this pupil dilation increases with increasing distance from the attended stimuli to eye fixation. Importantly, the

[2] An experiment that addresses this question is currently being conducted. The methods and materials, as well as some preliminary data, can be accessed at https://osf.io/grzup/.

differences in pupil size occurred independently of a pupillary light reflex. We believe that the presented results will be useful not only for researchers who investigate particular properties of the attentional system, but also for researchers who investigate visual shifts of attention in real-life situations, such as during driving or in team sports.

Appendix: Model syntax and comparisons for the data of the three experiments are presented below

For both the behavioral and the pupil size data, we first fitted a null model that only included the intercepts for participants and items. We then successively increased model complexity by adding an additional predictor, the linear combination of two predictors, or the interaction of two predictors (depending on the deign of the experiment). When a more complex model better fit the data in a single comparison than a less complex model, then the more complex model was used as basis for comparison with the next more complex model. For all model comparisons, we used log-likelihood ratio tests.

Final models for Experiment 1

Response accuracy

Model syntax	Chisq	df	*p*
response ~ 1 + (angle |participant) + (1|item)			
response ~ angle + (angle |participant) + (1|item)	55.5	4	<.001

Pupil size deviation from baseline

Model syntax	Chisq	df	*p*
pupil size ~ 1 + (1 |participant) + (1 |item)			
pupil size ~ angle + (1 |participant) + (1 |item)	72.8	4	<.001

angle = angle between stimuli and eye fixation (12.5°, 20°, 27.5°, 35°, or 42.5°)

Final models for Experiment 2

Response accuracy

Model syntax	Chisq	df	*p*
response ~ 1 + (angle |participant) + (1 |item)			
response ~ angle + (angle |participant) + (1 |item)	3.5	2	.178
response ~ task + (angle |participant) + (1|item)	149.1	1	<.001
response ~ angle + task + (angle |participant) + (1 |item)	18.4	2	<.001
response ~ angle * task + (angle |participant) + (1 |item)	9.6	2	.008

Pupil size deviation from baseline

Model syntax	Chisq	df	*p*
pupil size ~ 1 + (angle + task |participant) + (1 |item)			
pupil size ~ angle + (angle + task |participant) + (1 |item)	10.4	2	.005
pupil size ~ angle + task + (angle + task |participant) + (1 |item)	23.9	1	<.001
pupil size ~ angle * task + (angle + task |participant) + (1 |item)	2.9	2	.231

angle = angle between stimuli and eye fixation (12.5°, 27.5°, or 42.5°); *contrast* = stimuli presented in high contrast or low contrast (see main text)

Final models for Experiment 3

Response accuracy

Model syntax	Chisq	df	*p*
response ~ 1 + (contrast |participant) + (1 |item)			
response ~ contrast + (contrast |participant) + (1 |item)	8.5	1	.004
response ~ angle + contrast + (contrast |participant) + (1 |item)	53.1	2	<.001
response ~ angle * contrast + (contrast |participant) + (1 |item)	0.7	2	.717

Pupil size deviation from baseline

Model syntax	Chisq	df	*p*
pupil size ~ 1 + (contrast |participant) + (1 |item)			
pupil size ~ contrast + (contrast |participant) + (1 |item)	3.1	1	.079
pupil size ~ angle + (contrast |participant) + (|item)	5.2	2	.076
pupil size ~ angle + contrast + (contrast |participant) + (1 |item)	8.3	3	.041
pupil size ~ angle * contrast + (contrast |participant) + (1 |item)	6.1	2	.048

angle = angle between stimuli and eye fixation (12.5°, 27.5°, or 42.5°); *contrast* = stimuli presented in high contrast or low contrast (see main text)

References

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*, 276–292.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinari, G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 142–162). Cambridge, UK: Cambridge University Press.

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*, *33*, 2199–2204.

Binda, P., Pereverzeva, M., & Murray, S. O. (2014). Pupil size reflects the focus of feature-based attention. *Journal of Neurophysiology*, *112*, 3046–3052.

Blom, T., Mathôt, S., Olivers, C. N. L., & Van der Stigchel, S. (2016). The pupillary light response reflects encoding, but not maintenance, in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1716–1723. https://doi.org/10.1037/xhp0000252

Brocher, A., & Graf, T. (2016). Pupil old/new effects reflect stimulus encoding and decoding in short-term memory. *Psychophysiology*, *53*, 1823–1835.

Brocher, A., & Graf, T. (2017). Decision-related factors in pupil old/new effects: Attention, response execution, and false memory. *Neuropsychologia*, *102*, 124–134.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525

Crawford, B. H. (1936). The dependence of pupil size upon external light stimulus under static and variable conditions. *Proceedings of the Royal Society: Series B*, *121*, 376–395.

Gabay, S., Pertzov, Y., & Henik, A. (2011). Orienting of attention, pupil size, and the norepinephrine system. *Attention, Perception, & Psychophysics*, *73*, 123–129.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*, 457–461.

Granholm, E., & Verney, S. P. (2004). Pupillary responses and attentional allocation problems on the backward masking task in schizophrenia. *International Journal of Psychophysiology*, *52*, 37–51.

Heaver, B., & Hutton, S. B. (2011). Keeping an eye on the truth? Pupil size changes associated with recognition memory. *Memory*, *19*, 398–405.

Hüttermann, S., & Memmert, D. (2015). The influence of motivational and mood states on visual attention: A quantification of systematic differences and casual changes in subjects' focus of attention. *Cognition and Emotion*, *29*, 471–483.

Hüttermann, S., & Memmert, D. (2017). The attention window: A narrative review of limitations and opportunities influencing the focus of attention. *Research Quarterly for Exercise and Sport*, *88*, 169–183.

Hüttermann, S., Memmert, D., & Simons, D. J. (2014). The size and shape of the attentional "spotlight" varies with differences in sports expertise. *Journal of Experimental Psychology: Applied*, *20*, 147–157.

Hüttermann, S., Memmert, D., Simons, D. J., & Bock, O. (2013). Fixation strategy influences the ability to focus attention on two spatially separate objects. *PLoS ONE*, *8*, e65673.

James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt.

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly Journal of Experimental Psychology*, *64*, 1971–1989.

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*, 3080–3093.

Kafkas, A., & Montaldi, D. (2015). The pupillary response discriminates between subjective and objective familiarity and novelty. *Psychophysiology*, *52*, 1305–1316.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

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*, 175–185.

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), 1–9. https://doi.org/10.1167/14.14.7

Mathôt, S., Melmi, J.-B., van der Linden, L., & Van der Stigchel, S. (2016). The mind-writing pupil: A human–computer interface based on decoding of covert attention through pupillometry. *PLoS ONE*, *11*, e0148805. https://doi.org/10.1371/journal.pone.0148805

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

Mathôt, S., & Van der Stigchel, S. (2015). New light on the mind's eye: The pupillary light response as active vision. *Current Directions in Psychological Science*, *24*, 374–378

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology*, *4*, 919. https://doi.org/10.3389/fpsyg.2013.00919

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, *83*, 56–64.

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*, 560–569.

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences*, *91*, 7398–7403.

Richer, F., & Beatty, J. (1985), Pupillary dilations in movement preparation and execution. *Psychophysiology*, *22*, 204–207.

Richer, F., Silverman, C., & Beatty, J. (1983). Response selection and initiation in speeded reactions: A pupillometric analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 360–370.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*, 679–692.

Unsworth, N., & Robison, M. K. (2017). Pupillary correlates of covert shifts of attention during working memory maintenance. *Attention, Perception, & Psychophysics*, *79*, 782–795. https://doi.org/10.3758/s13414-016-1272-7

Van Gerven, P. W. M., Paas, F., Van Merrienboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*, 167–174.

Wahn, B., Ferris, D. P., Hairston, W. D., & König, P. (2016). Pupil sizes scale with attentional load and task experience in a multiple object tracking task. *PLoS ONE*, *11*, e168087:1–15. https://doi.org/10.1371/journal.pone.0168087

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, *51*, 277–284.