RESEARCH ARTICLE

Pupil Sizes Scale with Attentional Load and Task Experience in a Multiple Object Tracking Task

Basil Wahn[1]*, Daniel P. Ferris[2], W. David Hairston[3], Peter König[1,4]

1 Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany, **2** Human Neuromechanics Laboratory, School of Kinesiology, University of Michigan – Ann Arbor, MI, United States of America, **3** Human Research and Engineering Directorate, U.S. Army Research Laboratory, Aberdeen, MD, United States of America, **4** Department of Neurophysiology and Pathophysiology, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

* bwahn@uos.de

Abstract

Previous studies have related changes in attentional load to pupil size modulations. However, studies relating changes in attentional load and task experience on a finer scale to pupil size modulations are scarce. Here, we investigated how these changes affect pupil sizes. To manipulate attentional load, participants covertly tracked between zero and five objects among several randomly moving objects on a computer screen. To investigate effects of task experience, the experiment was conducted on three consecutive days. We found that pupil sizes increased with each increment in attentional load. Across days, we found systematic pupil size reductions. We compared the model fit for predicting pupil size modulations using attentional load, task experience, and task performance as predictors. We found that a model which included attentional load and task experience as predictors had the best model fit while adding performance as a predictor to this model reduced the overall model fit. Overall, results suggest that pupillometry provides a viable metric for precisely assessing attentional load and task experience in visuospatial tasks.

Introduction

Typically, pupil size changes are linked to changes in the amount of light that reaches the retina (a reflex known as the "pupil light reflex" [1, 2]) or whether a close or distant object is viewed (a response known as the "pupil near response" [1, 3]). However, for about 120 years, pupil sizes have been found to be additionally and systematically influenced by cognitive processing [3, 4]. For instance, pupil size changes have been studied in decision-making [5–8], attention [4, 9–22], emotions [23, 24], language [25], and memory [26–29].

A structure in the brain that is functionally involved in the control of pupil size changes is the locus coeruleus [16, 30–33]. The locus coeruleus projects to a wide range of brain regions but primarily to regions related to attentional processing such as the parietal cortex and the superior colliculus [34, 35], suggesting that pupil sizes are influenced by attentional processing.

Relatedly, pupil size changes often have been associated with attentional processing when task difficulties were varied [9, 16, 18, 36–38]. For instance, when performing arithmetic tasks of varying difficulty, pupil sizes increase with increasing task difficulty [36–38] and these findings also generalize to other attention demanding tasks [9, 16, 18]. However, studies that have investigated increases in attentional processing over a larger spectrum of difficulty levels and on a finer scale are scarce. In fact, attentional demands typically range along two levels (e.g., [9, 18]).

A task that has been suggested to be ideal to study increases in attentional processing over a larger spectrum of difficulty levels [18, 39, 40] and associated physiological correlates [41] is the so-called multiple object tracking ("MOT") task [42]. In this task, a subset of several objects are indicated as "targets" at the beginning of a trial on a computer screen. Then, objects become indistinguishable and move randomly across the screen. After several seconds, objects stop moving and participants need to indicate which objects were the targets. In a recent study, Alnæs and colleagues [16] investigated how modulations in pupil sizes systematically changed as a function of attentional load in a MOT task (i.e., the number of targets that needed to be tracked). In particular, they varied the number of targets between two and five objects and also included a passive viewing condition in which participants did not track any targets. They found that pupil sizes increased with each increase in attentional load, suggesting that pupil size modulations are related to task difficulty also over a larger spectrum of attentional load manipulations. The stability of the pupil size modulations by the attentional load in the MOT task was assessed by correlating the modulations in a first session with modulations in several subsequent recording sessions that were recorded within the span of several years. They found a high correlation between sessions, suggesting that pupil size modulations related to attentional load are stable over a long time. Overall, the study by Alnæs and colleagues [16] found a close relationship between pupil sizes and attentional load in a MOT task that is stable over a long time.

These insights [16], however, raise a number of new questions. An important aspect that has not been addressed is the influence of task experience on pupil size modulations. In particular, it was not investigated how training effects over several sessions within a few days can modulate pupil sizes. Repeated performance of tasks such as a MOT task leads to improved performance. Thus, the repeated execution of a MOT task could lead to a dissociation between attentional load and task difficulty (i.e., the attentional load manipulation remains constant while the task difficulty decreases with task experience). We hypothesized that pupil sizes could systematically decrease with increasing task experience as participants adapt to the task requirements and devote less processing resources to the task. Such a finding would demonstrate that pupil size modulations are also affected by task experience and not merely by the attentional load manipulations in the MOT task.

Another important question is how performing a task as such and increasing the attentional load in a task affect pupil sizes. In the study by Alnæs and colleagues [16], a condition in which participants tracked only one target was not included (i.e., tracking load was only varied between two and five targets, and a passive viewing condition was included). However, including a one target condition does allow to compare pupil size modulations relative to the passive viewing condition (i.e., when no MOT task is performed and the moving objects are passively viewed). We hypothesized that this comparison involves two factors that influence pupil size modulations: 1) a change in attentional load (i.e., tracking one target more) and 2) the switch in attention demands from performing no task to performing a task. For the remaining comparisons involving conditions in which targets are tracked (e.g., between tracking one and two targets), we hypothesized that only the former factor (i.e., the increase of attentional load) would affect pupil sizes. Therefore, we predicted that the magnitude of the pupil size

modulations between performing no task in comparison to performing a task should be larger than the pupil size modulations due to performing a task at a lower compared to a higher task difficulty level. As a point of note, in both these cases the attentional load manipulation is the same (i.e., the number of tracked targets is increased by one).

In addition, given all these partially redundant factors that could influence pupil size modulations (i.e., attentional load, task performance, and task experience), we address the question of which combination of this redundant set of predictors leads to the highest explanatory power (i.e., model fit) for predicting pupil size modulations.

Taken together, we address how changes in attentional load that span the complete spectrum of load manipulations (i.e., from no load manipulation up to a condition of high attentional load) systematically affect pupil sizes. Moreover, we address to what extent attentional load, task experience and/or performance influence pupil sizes.

Results

We investigated 20 participants' (14 female, $M = 23.03$ years, $SD = 3.34$ years) performance in a MOT task at varying levels of attentional load (Fig 1). In particular, the number of targets to be tracked varied from zero, which served as a passive viewing condition, up to five targets. This allowed us to investigate whether pupil size changes between a passive viewing condition and tracking one target are different from changes in pupil sizes between conditions in which targets are tracked. Moreover, to investigate effects of task experience, each participant performed the experiment on three consecutive days.

An overview of the MOT task performance (i.e., fraction of correctly selected targets) is shown in Fig 2A. Performance systematically decreased with an increasing number of targets in the MOT task. Moreover, participants slightly improved their tracking performance over days. We tested whether these observations were statistically reliable using linear mixed models and model comparisons [43]. For this purpose, we first constructed a baseline model, in which only intercepts were fitted as a random effect (i.e., intercepts were fitted for each participant individually to account for the repeated measures design [44]). Then, we subsequently added predictors to the baseline model and assessed via model comparisons whether these added predictors significantly explain additional variance in the data. As a model selection
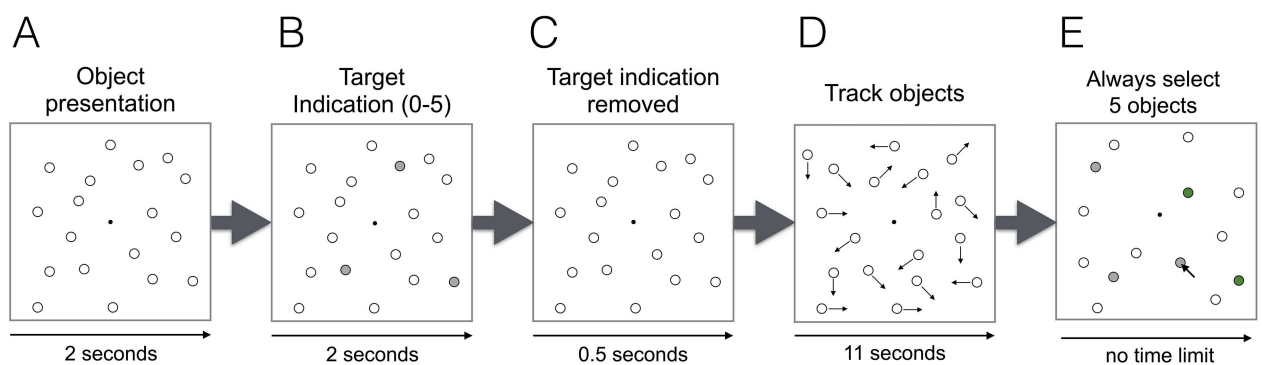

Fig 1. Multiple object tracking (MOT) task trial overview. (A) Objects (total: 18) are shown. (B) A subset of objects (i.e., the "targets", varied between zero and five) turn gray for two seconds. Note that in the passive viewing condition (zero targets), no objects turn gray in this phase. (C) The target indication is removed. Targets are now indistinguishable from the other objects. (D) Objects randomly move across the screen and participants need to track the movements of the targets. Arrows indicate the current movement direction of the objects. (E) The objects stop moving and participants select the targets using the mouse. To keep motor demands constant, participants always need to select five objects but only the first selections count to performance (e.g., if two targets need to be tracked, then only the first two selections count to performance). After five objects are selected, feedback is given (i.e., correct selections turn green).

doi:10.1371/journal.pone.0168087.g001

criterion, we report the Bayesian Information Criterion ("BIC") for each addition. The BIC is a measure of how well a model fits data that is corrected for the number of included predictors. We first added attentional load (i.e., the number of targets tracked in the MOT task) to the model as a fixed and random effect (i.e., slopes are fitted for each participant individually) and found that these additions significantly explained additional variance in the data ($\chi^2(3) = 170.31$, $p < .001$, BIC = 2195.5—for comparison, the baseline model's BIC = 2348.7). We then added task experience (i.e., the index of day of measurement) as a fixed and random effect to the model and found that these additions significantly explain additional variance in the data ($\chi^2(4) = 38.13$, $p < .001$, BIC = 2180.2). Adding an interaction effect between attentional load and task experience did not significantly explain additional variance in the data ($\chi^2(1) = 1.17$, $p = .280$, BIC = 2184.8).

For the final model, in which only attentional load and task experience were included, we assessed the regression coefficients and their significance. We found a significant negative slope ($B = -4.67$, SE = 0.62, $p < .001$) for attentional load, suggesting that participants' performance on average decreased with increasing number of targets by about five percent for each added target. The regression coefficient for task experience was not significant ($B = 1.62$, SE = 0.98, $p = .114$). However, as noted above, using a more sensitive test, the model comparisons when adding task experience to the model yielded a significant comparison, suggesting that task experience does explain variance in task performance in addition to the variance explained by attentional load.

In sum, participants' performance in the MOT task decreased with increasing attentional load and we found a slight improvement in performance with increasing task experience. These findings indicate that the experimental manipulation of attentional load (i.e., varying the number of tracked targets) was successful in having the intended effect as task performance systematically decreased with increasing attentional load.

As a next step, we investigated pupil sizes (Fig 2B, 2C and 2D) and their relation to attentional load, task performance and task experience. We averaged over the median pupil sizes of the left and right eye and normalized the resulting values relative to the passive viewing condition (for more details see methods section "Methods of data analysis"). Note that although the total moving time amounted to eleven seconds, pupil size data were only extracted between three and nine seconds in which the objects were moving (a similar time window was also used in [16]). This circumvents potential perceptual effects like contrast changes during target indication and executive effects like motor preparation to select targets after objects' motion in the data. In addition, when inspecting pupil size traces across the whole eleven seconds (see S1 Fig), pupil size differences between attentional load conditions appear to develop within the first seconds of a trial and then remain relatively constant throughout the rest of the trial.

Pupil sizes systematically increased with attentional load (Fig 2B) and decreased with task experience (Fig 2C). When relating pupil sizes to the MOT performance, a highly negative relation can be observed (i.e., with increasing performance, pupil sizes decrease) (Fig 2D). Summarizing, on a descriptive level, the data suggests that pupil size modulations depend on the attentional load, task experience, and task performance. Moreover, these observations suggest that this set of variables is partially redundant in explaining variance in the pupil size modulations.

On an inferential level, to investigate which are the most relevant variables in the redundant set (i.e., attentional load, task experience, and task performance), we used linear mixed models and model comparisons in close analogy to the investigation of MOT task performance above. In the baseline model, we fitted intercepts for each participant individually to account for the repeated measurements design. We then added attentional load as a fixed and random effect to the model. We found that in comparison to the baseline model (BIC = 1611.1, Fig 3 box "Base
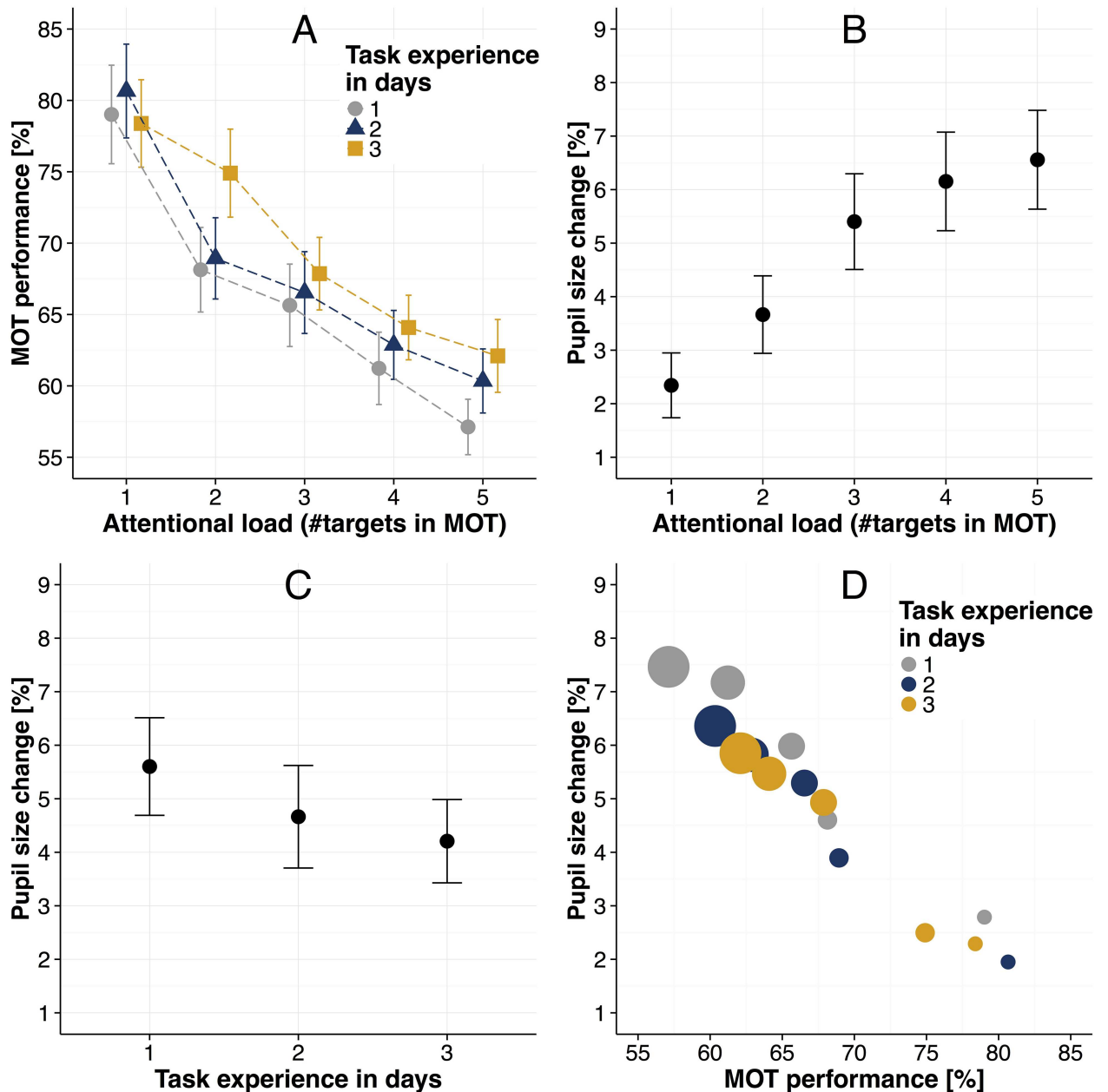
Fig 2. Main experiment results. (A) MOT task performance (i.e., fraction of correctly selected targets) as a function of attentional load and task experience. (B) Pupil size changes relative to the passive viewing condition as a function of attentional load. (C) Pupil size changes relative to the passive viewing condition as a function of task experience. (D) MOT task performance plotted against the pupil size changes, separately for each day. In addition, the number of targets in the MOT task is coded by the sizes of the circles (i.e., with increasing circle size, the number of targets increases). Error bars in all panels are standard error of the mean.

doi:10.1371/journal.pone.0168087.g002

model") these additions significantly explained more variance in the data ($\chi^2(3) = 96.65$, $p < .001$, BIC = 1531.6, Fig 3 box "Attentional load"). We then added task experience as an additional fixed and random effect to the model and found that these additions were significant as well ($\chi^2(4)=40.59$, $p < .001$, BIC = 1513.8, Fig 3 box "Task experience, Attentional load"). As a point of note, when inspecting the pupil size traces across the tracking period for each day (see S1 Fig), the effect of task experience appears to be driven by a greater reduction
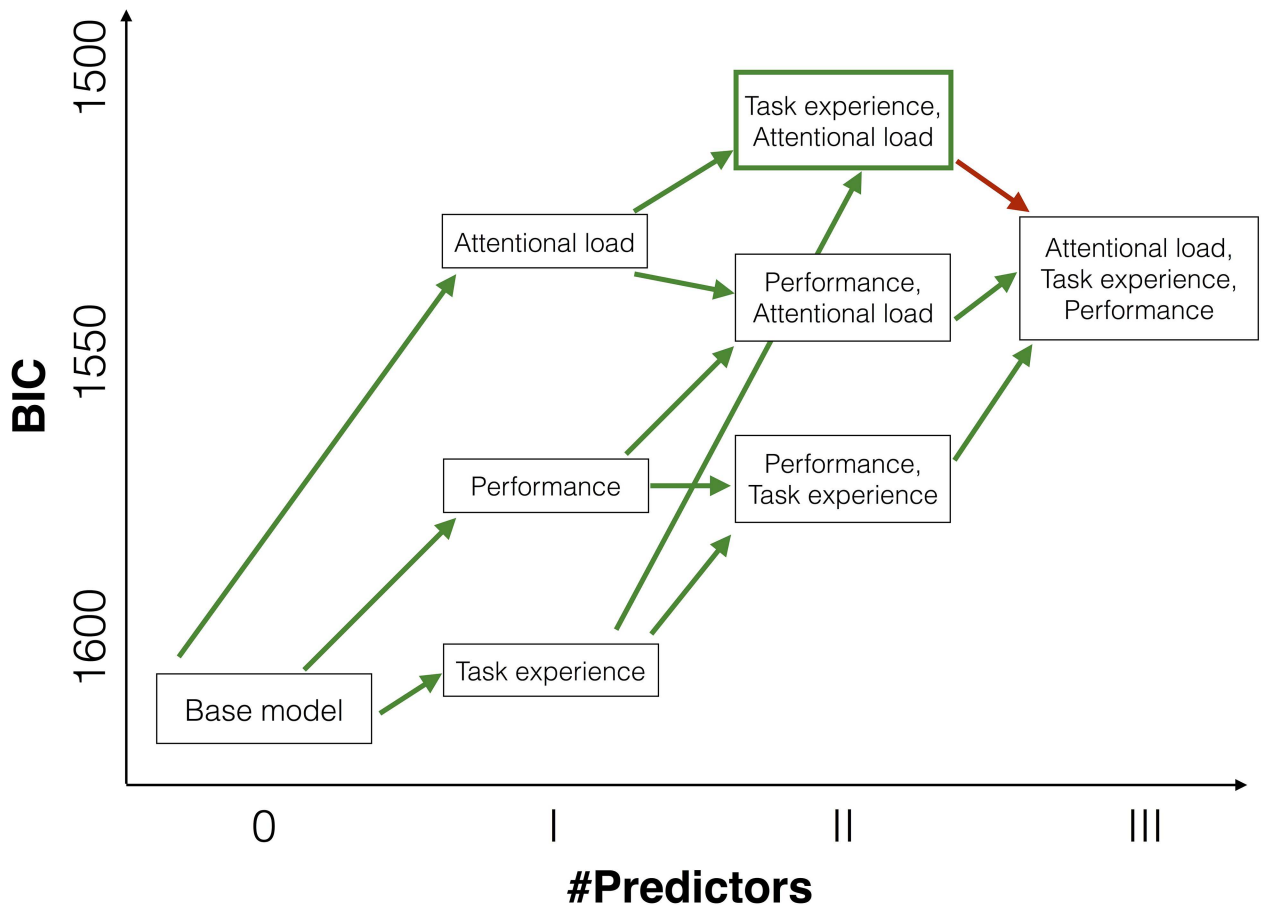
Fig 3. Model comparisons overview. The ordinate shows the model fit in form of the reported Bayesian information criterion (BIC). The abscissa indicates the number of predictors in the model. Green arrows indicate model additions that significantly explained additional variance to the variance explained by previous predictors. Red arrows indicate that the added predictors did not explain additional variance.

of pupil sizes (over days) after an initial increase in pupil sizes in the first few seconds of a trial. Adding an interaction effect between task experience and attentional load as a fixed and random effect to the model did not lead to a significant increase in explained variance ($\chi^2(5)$=5.25, p = .386, BIC = 1537.0), suggesting that there is no interaction effect between task experience and attentional load.

In principle, attentional load and task experience might already sufficiently predict changes in the pupil sizes while performance in the MOT task does not explain substantial additional variance. Conversely, it could be that the task performance in the MOT task explains significant additional variance in the pupil size data and replaces none, one, or both of the other variables. To investigate this question, we extended the modeling approach described above. We used attentional load, task experience, and task performance as predictors in all possible combinations to predict pupil size modulations. Then, we computed whether adding a variable does significantly explain more variance in the data (Fig 3). Notably, for the model that had attentional load and task experience as predictors, we found that adding performance as an additional predictor did not explain additional variance in the data ($\chi^2(5)$=8.16, p = .150, BIC = 1534.2, see red arrow in Fig 3). Thus, the complete model including all three predictors was not the best. Yet, the performance could significantly explain additional variance if it was

added to a model in which only the predictor attentional load ($\chi^2(4) = 14.10$, $p = .007$, BIC = 1540.3) or task experience ($\chi^2(4)=63.89$, $p < .001$, BIC = 1567.1) was included. Moreover, of all possible models, when comparing the BICs, a model considered above with attentional load and task experience, but not including task performance was preferred (BIC = 1513.8). Overall, we conclude that task performance did not explain enough additional variance on top of the variance explained by attentional load and task experience in the pupil size data in order to justify its inclusion. Moreover, given the BICs for all tested models, a model including only attentional load and task experience as predictors is preferred.

We then assessed the regression coefficients and their significance for the model that provided the best BIC (i.e., the model including attentional load and task experience as predictors). We found a significant positive slope for attentional load ($B = 1.09$, $SE = 0.13$, $p < .001$), suggesting that participants' pupil sizes increased by about one percent per added target. Moreover, we found a significant negative slope for task experience ($B = -0.70$, $SE = 0.32$, $p = .04$), indicating that participants' pupil sizes decreased with increasing experience by about 0.7% per day.

As a point of note, with regard to the finding that pupil sizes systematically scale with attentional load in the MOT task, one might argue that increases in pupil sizes could be simply due to attending increasingly more dark edges with increasing targets in the MOT task and not due to increases in attentional load [13, 17, 45]. We investigated whether this potential confound applies to our data by conducting a control experiment in which targets with bright edges were tracked on a dark background with a separate set of participants (N = 20, 16 female, $M = 21.70$ years, $SD = 3.31$ years) that were measured for only one day. Except for changing the color of the background and edges of the targets, the experiment was exactly the same (see Fig 4A and 4B for the MOT performance and pupil sizes). For this set of participants, we again found that adding attentional load as a predictor to the baseline model significantly explained more variance in the data ($\chi^2(2) = 22.20$, $p < .001$, BIC = 516.4, baseline model's BIC: 529.4). We then assessed the regression coefficient and it's significance for this model and again found a significant positive slope for attentional load ($B = 0.87$, $SE = 0.22$, $p = .001$), suggesting that also for this control experiment participants' pupil sizes increased by about one percent per added target.

Regarding the MOT performance of the control experiment, we also replicated the results of the main experiment. That is, when adding the predictor attentional load as a fixed and random effect to a baseline model, we found that these additions significantly explain additional variance in the data ($\chi^2(2) = 26.14$ $p < .001$, BIC = 739.5, baseline model's BIC: 756.5). Assessing the regression coefficient for the predictor attentional load again yielded a significant negative slope of similar magnitude as in the main experiment ($B = -4.14$, $SE = 0.78$, $p < .001$), suggesting that for the control experiment the attentional load manipulation also had the desired effect.

As a next step, we compared whether pupil size changes between tracking one target and the passive viewing condition (i.e., the difference between performing the MOT and no task at all) differ from the changes between tracking conditions (e.g., the difference between tracking one and two targets). For this purpose, we extracted the fitted slopes (i.e., a slope is fitted for each participant separately) from the final model (i.e., the model including attentional load and task experience as predictors) and the intercepts for each participant. We compared the extracted slopes and intercepts using a paired t-test. Note, the fitted intercepts in this model represent the difference between the attentional load condition with one target and the passive viewing condition while the slopes represent the average increase in pupil sizes when the attentional load is increased by one unit, i.e., one target. We found that intercepts are significantly larger than slopes (mean difference: 1.85%, $t(19) = 2.78$, $p = .012$), suggesting that the
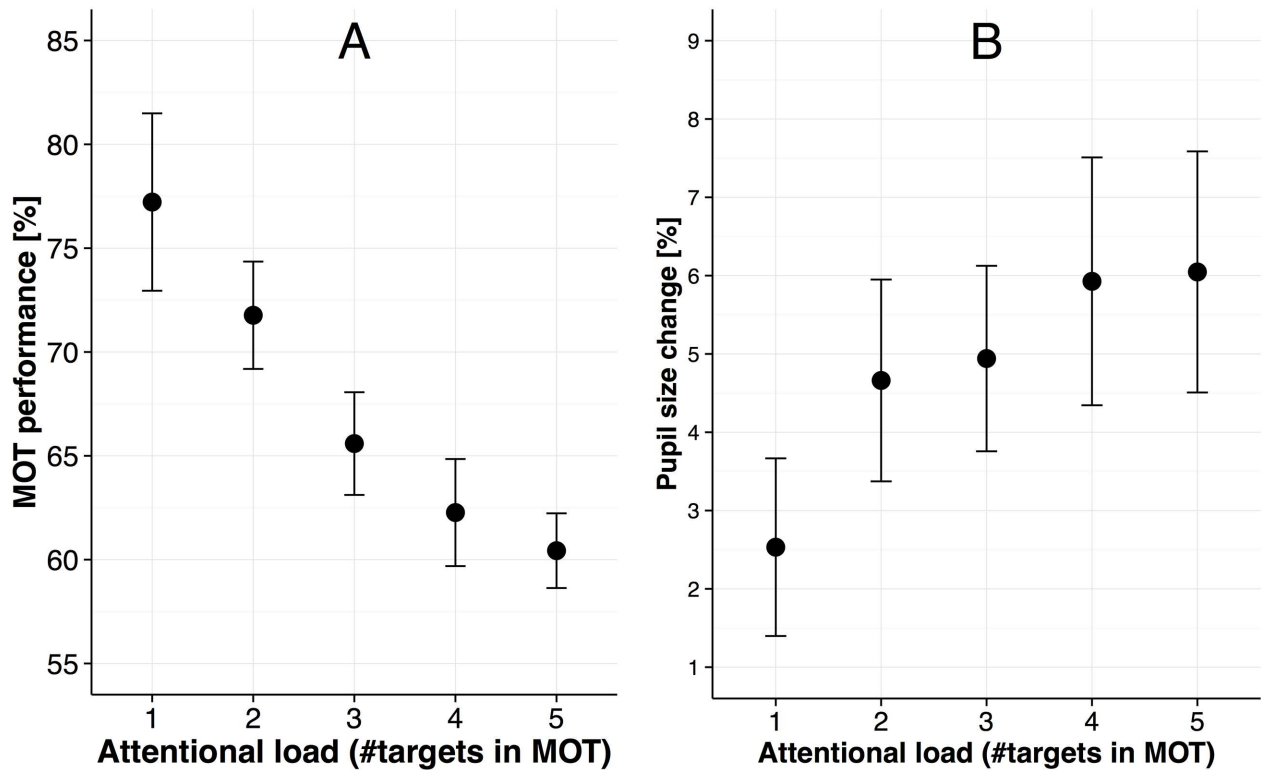
Fig 4. Control experiment results. (A) MOT task performance (i.e., fraction of correctly selected targets) as a function of attentional load and task experience. (B) Pupil size change relative to the passive viewing condition as a function of attentional load. Error bars in all panels are standard error of the mean.

doi:10.1371/journal.pone.0168087.g004

difference in pupil sizes between the passive viewing condition and tracking one target condition is about 70% larger than the average difference between tracking conditions. We also tested whether this effect persists when we restricted the data to attentional load conditions one to four, one to three, and one to two, respectively. While on a descriptive level the direction of the effect was still present, we found that this effect was considerably smaller and no longer significant for these subsets of the data (one to two: 1.00% difference, $t(19) = 1.56$, $p = .135$; one to three: 0.43% difference, $t(19) = 0.86$, $p = .403$; one to four: 1.12%, $t(19) = 1.96$, $p = .065$). These findings suggest that the effect above for all attentional load conditions is potentially modulated by a ceiling effect (i.e., for each increase in attentional load the pupil sizes tend to become smaller).

Discussion

In line with the results of the study by Alnæs and colleagues [16], we found that pupil sizes and tracking performance were systematically influenced by attentional load in a MOT task (i.e., the number of tracked targets). These findings suggest that pupil sizes reflect the current deployment of attentional resources to a task also on a finer scale as each increase in attentional load was accompanied by an increase in pupil sizes.

What is more, we found that repeating the experiment on two consecutive days led to a significant decrease in pupil sizes independent of the number of tracked targets in the MOT task. Moreover, these pupil size modulations over days were accompanied by slight performance improvements. Relatedly, previous work has shown that repeated presentations of a stimulus

does lead to a pupil size decrease in a memory task [27], indicating a decrease in mental effort to retrieve the viewed stimulus from memory. Here, we show for the first time that performing an attentionally demanding task such as the MOT task on three consecutive days similarly led to not only an increase in performance, as might be expected, but also a concomitant decrease in pupil size as well. This reduction in pupil sizes over several days suggests a reduction of the attentional demand to perform the MOT task related to having extended experience with performing the task.

More generally, our findings suggest an interdependency between the performance in the MOT task and pupil sizes. Such an interplay may explain additional variance in the data beyond the variance explained by the attentional load and the task experience. We further investigated this observation by testing all possible models involving attentional load, task experience, and task performance as predictors. We found that the performance in the MOT task did not substantially explain additional variance on top of what was already explained by the number of tracked targets and days of measurements. Contrasting earlier findings that suggested that pupil size differences also reflect interindividual differences in information processing [46, 47], these findings suggest that changes in pupil sizes are sufficiently explained by the level of attentional load and task experience.

In addition, we found a higher pupil size increase for the comparison involving performing no task and tracking only one target than for the average pupil size increases for each additionally tracked target. This result suggests that pupil sizes not only reflect the current deployment of attentional resources related to attentional load but additionally reflect whether subjects are engaged in a task or not. However, we also found that this effect was considerably reduced in size and no longer significant when we restricted our analysis to a smaller set of attentional load conditions. Such a finding suggests that this effect may be modulated by a ceiling effect. That is, with increasing number of targets in the MOT task, the pupil size increases between attentional load conditions become systematically smaller. To resolve these issues, future studies could investigate attentional load conditions that are not prone to ceiling effects with a higher statistical power.

In conclusion, we find that pupil sizes are modulated by attentional load on a fine scale, i.e. each increment in attentional load led to a further increase in pupil sizes. In addition, pupil sizes are additionally modulated by whether humans are actively engaged in a task or not. Furthermore, pupil size modulations are systematically reduced with increasing task experience and improving task performance, suggesting also a close link between task performance and pupil sizes on a group level. Yet, modeling individual participants' pupil sizes task performance does not additionally explain variance on top of the variance explained by the manipulations of attentional load and task experience. In fact, a model incorporating only the two predictors attentional load and task experience is selected as the best. Overall, this study demonstrates that pupil size modulations are affected by a wide range of different factors related to performing a task. These factors range from whether humans are actually performing a task or not, changes in attentional load, and changes in task experience.

More generally, the present findings suggest that pupil sizes could provide a general metric to assess attentional load in attention demanding visuospatial tasks (e.g., in visual search [48, 49] or visuomotor tasks [50]) without necessarily taking the task performance into account. Findings could also be applicable to demanding real world tasks (e.g., air-traffic control, driving a car, or flying an airplane) to assess the current attentional load during task performance. In particular, given that pupil sizes are modulated by attentional load on such a fine scale consistently over days, critical levels of attentional load (i.e., task-relevant information is no longer effectively processed or neglected) could be detected during task performance and timely interventions could prevent potential accidents (also see [29, 51–55]).

Methods

Methods of data acquisition

Participants. All participants were students of the University of Osnabrück. 20 students (14 female, M = 23.03 years, SD = 3.34 years) participated in the main experiment. In a control experiment, a separate set of 20 students (16 female, M = 21.70 years, SD = 3.31 years) participated. The study was approved by the ethics committee of the University Osnabrück. We informed participants about their rights and all participants signed a written consent form. Participants either received a monetary reward or course credits for participation.

Experimental setup. Participants sat in a dark room in front of a computer screen (BenQ XL2420T, resolution 1920 × 1080, 120 Hz, subtending a visual field of 32.87 × 18.49 visual degrees) at a distance of 90 cm. Eye movements were recorded using an Eyelink II (binocular tracking, 500 Hz sampling rate). In order to calibrate eye position, we used a 5-point grid and repeated calibration until the mean error was below 0.7 degrees for both eyes. As pupil size detection procedure, we selected the "diameter" option.

Experimental procedure. In the experiment, participants were instructed to track a subset of randomly chosen objects ('targets') among eighteen randomly moving objects (for a task overview, see Fig 1). Objects were 1.06 visual degrees in width. The subset of randomly chosen targets was varied between zero (a "passive viewing" condition) up to a maximum of five targets. At the beginning of a trial, all objects were shown for two seconds. Then, a subset of objects turned gray for two seconds, indicating the targets and then reverted back to look like all the other objects. After an additional 500 ms, objects started to move and participants were instructed to track the targets. While objects were moving, they repelled each other and bounced off the screen borders. The objects' movement direction and velocity was randomly chosen within each frame with a probability of 0.01. With a refresh rate of 100 Hz this resulted in an average velocity of 2.57 visual degrees per second [minimum 1.71, maximum 3.42]. Objects were moving for a total of eleven seconds. During objects' motion, all objects looked the same. After objects' motion seized, participants were instructed to select the targets with a computer mouse. We suspected that preparing for selecting a certain number of targets could systematically influence pupil sizes across attentional load conditions (see supporting information in [56]). Therefore, to keep this executive processing demand constant in all trials, participants were instructed to always select five objects regardless of the number of targets that they were required to track. Importantly, participants were instructed to always select those objects that they think are the targets first. Participants were also informed that only the targets that were selected first counted towards their performance and only the number of selections matching the number of targets (e.g., if participants need to track two targets, then only the first two target selections count). After five objects were selected, correctly selected objects were shown in green. Participants were instructed to fixate on the fixation point (0.15 visual degrees radius) in the center of the screen while the objects moved randomly across the screen. Before each trial, a drift correction was applied for which participants needed to fixate the fixation point and press the space key on the keyboard. To ensure that the perceptual load is kept constant, participants always saw a total of eighteen objects in each trial. Before the experiment started, participants were familiarized with the task with three training trials that always contained two, three and zero targets. In the experiment, participants completed a total of 120 trials. Trials were grouped in blocks of six trials. In each block, each number of targets (i.e., zero to five) was done once in a randomized order. Repetition of the same target number in consecutive trials was avoided. The complete experiment took approximately one hour per participant.

Given the relatively large amount of trials per day, we assessed whether any general fatigue effects occurred by correlating the MOT performance with the number of trials for each day and participant separately. Generally, correlations were small and positive (mean r = .02), suggesting that there were no fatigue effects. However, correlations were significantly different from zero for the first (mean r = .04, t = 2.30, p = .033) and second (mean r = .04, t = 2.16, p = .043) but not for the third day (mean r = .01, t = 0.39, p = .702) of the measurements. This suggests that small training effects also occurred within an experimental session for the first two days of the measurements.

Methods of data analysis

We downsampled the eye tracking data to 100 Hz. Data points with physiologically unlikely pupil sizes (i.e., negative pupil sizes as produced by blinks) within a trial were removed for each eye separately. Moreover, for each eye, we removed outliers within a trial using boxplots and applied 3 times the interquartile range as a threshold. Trials in which participant's gaze deviated from the fixation point in the center by more than three visual degrees on average were removed from the analysis (for the main experiment: total of 8.38% trials excluded, M = 7.62 visual degrees, SD = 5.05 visual degrees; for the control experiment: total of 8.32% trials excluded, M = 10.66 visual degrees, SD = 9.37 visual degrees). Finally, for each trial and for each eye, we took the median pupil size for the time interval between three and nine seconds of the tracking period (i.e., when objects randomly moved across the screen). We chose this time interval to avoid early sensory confounds (i.e., contrast changes due to seeing a variable number of targets prior to the onset of the object motion) within the first several seconds in a trial and late executive confounds (i.e., effects of preparing motor actions to select targets at the end of a trial). A similar time interval has also been used in the study by Alnæs and colleagues [16]. We chose the median as a measure of central tendency as it is more robust towards outliers in the data. Specifically, these outliers could be implausible data points that are not filtered out by the boxplots that might occur due to technical malfunctions or remaining data points that preceded or followed blinks.

For investigating pupil size changes, we averaged across the median pupil sizes of the left and right eye for each trial. For the purpose of normalizing the data relative to the passive condition, all trials were averaged for each participant separately for each number of targets in the MOT task and day. Resulting values for targets between one and five were normalized relative to the passive viewing condition. In particular, we calculated the percent increase in pupil sizes relative to the passive viewing condition by dividing the pupil sizes in the target conditions by the pupil sizes in the passive viewing condition. As the attentional load conditions were presented in a randomized order (see section "Experimental Procedure" above for more details) the normalization step was not reasonable to perform on the trial level. Therefore, we performed the normalization step on the aggregated data. Moreover, due to randomly interleaving attentional load conditions, potential accumulations of pupil size increases over trials were averaged out.

Supporting Information

S1 Fig. Figure of pupil size traces as a function of attentional load for each day condition. (TIFF)

S1 File. Zip file containing aggregated datasets underlying Figs 2, 3 and 4. (ZIP)

S2 File. Link to the raw MOT task performance and eyetracking data (osf.io/qtzjb).
(PDF)

Acknowledgments

Author Contributions

Conceptualization: BW PK DPF WDH.

Data curation: BW.

Formal analysis: BW.

Funding acquisition: PK.

Investigation: BW.

Methodology: BW PK.

Resources: BW PK.

Software: BW.

Supervision: PK.

Visualization: BW.

Writing – original draft: BW.

Writing – review & editing: BW PK DPF WDH.

References

1. Loewenfeld IE. Mechanisms of reflex dilatation of the pupil. Documenta Ophthalmologica. 1958; 12(1): 185–448. doi: 10.1007/BF00913471 PMID: 13609524

2. Sirois S, Brisson J. Pupillometry. Wiley Interdisciplinary Reviews: Cognitive Science. 2014; 5(6): 679–692. PMID: 26308873

3. Beatty J, Lucero-Wagoner B. The pupillary system. Handbook of Psychophysiology. 2000; 2:142–162.

4. Heinrich W. Die Aufmerksamkeit und die Funktion der Sinnesorgane. Zeitschrift für Psychologie und Physiologie der Sinnesorgane. 1896; p. 342–388.

5. Einhäuser W, Koch C, Carter O. Pupil dilation betrays the timing of decisions. Frontiers in human neuroscience. 2010; 4:18. doi: 10.3389/fnhum.2010.00018 PMID: 20204145

6. Preuschoff K, Marius't Hart B, Einhäuser W. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. Frontiers in neuroscience. 2011; 5. doi: 10.3389/fnins.2011.00115 PMID: 21994487

7. de Gee JW, Knapen T, Donner TH. Decision-related pupil dilation reflects upcoming choice and individual bias. Proceedings of the National Academy of Sciences. 2014; 111(5):E618–E625. doi: 10.1073/pnas.1317557111

8. Murphy PR, Vandekerckhove J, Nieuwenhuis S. Pupil-linked arousal determines variability in perceptual decision making. PLOS Comput Biol. 2014; 10(9):e1003854. doi: 10.1371/journal.pcbi.1003854 PMID: 25232732

9. Porter G, Troscianko T, Gilchrist ID. Effort during visual search and counting: Insights from pupillometry. The Quarterly Journal of Experimental Psychology. 2007; 60(2):211–229. doi: 10.1080/17470210600673818 PMID: 17455055

10. Geng JJ, Blumenfeld Z, Tyson TL, Minzenberg MJ. Pupil diameter reflects uncertainty in attentional selection during visual search. Frontiers in Human Neuroscience. 2015; 9. doi: 10.3389/fnhum.2015.00435 PMID: 26300759

11. Zylberberg A, Oliva M, Sigman M. Pupil dilation: a fingerprint of temporal selection during the "attentional blink". Frontiers in Psychology. 2012; 3. doi: 10.3389/fpsyg.2012.00316 PMID: 22973248

12. Wierda SM, van Rijn H, Taatgen NA, Martens S. Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. Proceedings of the National Academy of Sciences. 2012; 109(22):8456–8460. doi: 10.1073/pnas.1201858109

13. Naber M, Alvarez GA, Nakayama K. Tracking the allocation of attention using human pupillary oscillations. Frontiers in Psychology. 2013; 4. doi: 10.3389/fpsyg.2013.00919 PMID: 24368904

14. Mathôt S, Van der Linden L, Grainger J, Vitu F. The pupillary light response reveals the focus of covert visual attention. PLoS ONE. 2013; 8(10):e78168. doi: 10.1371/journal.pone.0078168 PMID: 24205144

15. Mathôt S, Melmi JB, van der Linden L, Van der Stigchel S. The mind-writing pupil: A human-computer interface based on decoding of covert attention through pupillometry. PLOS ONE. 2016; 11(2):e0148805. doi: 10.1371/journal.pone.0148805 PMID: 26848745

16. Alnæs D, Sneve MH, Espeseth T, Endestad T, van de Pavert SHP, Laeng B. 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. 2014; 14(4):1. doi: 10.1167/14.4.1 PMID: 24692319

17. Binda P, Murray SO. Keeping a large-pupilled eye on high-level visual processing. Trends in Cognitive Sciences. 2015; 19(1):1–3. doi: 10.1016/j.tics.2014.11.002 PMID: 25467128

18. Wahn B, König P. Vision and Haptics Share Spatial Attentional Resources and Visuotactile Integration Is Not Affected by High Attentional Load. Multisensory Research. 2015; 28(3–4):371–392. doi: 10.1163/22134808-00002482 PMID: 26288905

19. Karatekin C, Couperus JW, Marcus DJ. Attention allocation in the dual-task paradigm as measured through behavioral and psychophysiological responses. Psychophysiology. 2004; 41(2):175–185. doi: 10.1111/j.1469-8986.2004.00147.x PMID: 15032983

20. Lisi M, Bonato M, Zorzi M. Pupil dilation reveals top–down attentional load during spatial monitoring. Biological Psychology. 2015; 112:39–45. doi: 10.1016/j.biopsycho.2015.10.002 PMID: 26472280

21. Hoeks B, Levelt WJ. Pupillary dilation as a measure of attention: A quantitative system analysis. Behavior Research Methods, Instruments, & Computers. 1993; 25(1):16–26. doi: 10.3758/BF03204445

22. Kloosterman NA, Meindertsma T, Loon AM, Lamme VA, Bonneh YS, Donner TH. Pupil size tracks perceptual content and surprise. European Journal of Neuroscience. 2015; 41(8):1068–1078. doi: 10.1111/ejn.12859 PMID: 25754528

23. Granholm E, Steinhauer SR. Pupillometric measures of cognitive and emotional processes. International Journal of Psychophysiology. 2004; 52(1):1–6. doi: 10.1016/j.ijpsycho.2003.12.001 PMID: 15003368

24. Vanderhasselt MA, Remue J, Ng KK, De Raedt R. The interplay between the anticipation and subsequent online processing of emotional stimuli as measured by pupillary dilatation: the role of cognitive reappraisal. Frontiers in Psychology. 2014; 5. doi: 10.3389/fpsyg.2014.00207 PMID: 24659976

25. Hyönä J, Tommola J, Alaja AM. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. The Quarterly Journal of Experimental Psychology. 1995; 48(3):598–612. doi: 10.1080/14640749508401407 PMID: 7568993

26. Kahneman D, Beatty J. Pupil diameter and load on memory. Science. 1966; 154(3756):1583–1585. doi: 10.1126/science.154.3756.1583 PMID: 5924930

27. van Rijn H, Dalenberg JR, Borst JP, Sprenger SA. Pupil dilation co-varies with memory strength of individual traces in a delayed response paired-associate task. PLoS ONE. 2012; 7(12):e51134. doi: 10.1371/journal.pone.0051134 PMID: 23227244

28. Johnson EL, Singley ATM, Peckham AD, Johnson SL, Bunge SA. Task-evoked pupillometry provides a window into the development of short-term memory capacity. Frontiers in Psychology. 2014; 5. doi: 10.3389/fpsyg.2014.00218

29. Causse M, Peysakhovich V, Fabre EF. High Working Memory Load Impairs Language Processing during a Simulated Piloting Task: An ERP and Pupillometry Study. Frontiers in Human Neuroscience. 2016; 10. doi: 10.3389/fnhum.2016.00240

30. Gilzenrat MS, Nieuwenhuis S, Jepma M, Cohen JD. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cognitive, Affective, & Behavioral Neuroscience. 2010; 10(2):252–269. doi: 10.3758/CABN.10.2.252

31. Wang CA, Munoz DP. A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. Current opinion in neurobiology. 2015; 33:134–140. doi: 10.1016/j.conb.2015.03.018 PMID: 25863645

32. Samuels E, Szabadi E. Functional neuroanatomy of the noradrenergic locus coeruleus: Its roles in the regulation of arousal and autonomic function part I: principles of functional organisation. Current Neuropharmacology. 2008; 6(3):235. doi: 10.2174/157015908785777229 PMID: 19506723

33. Rajkowski J, Majczynski H, Clayton E, Aston-Jones G. Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. Journal of Neurophysiology. 2004; 92(1):361–371. doi: 10.1152/jn.00673.2003 PMID: 15028743

34. Bouret S, Sara SJ. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends in Neurosciences. 2005; 28(11):574–582. doi: 10.1016/j.tins.2005.09.002 PMID: 16165227

35. Schneider KA, Kastner S. Effects of sustained spatial attention in the human lateral geniculate nucleus and superior colliculus. The Journal of Neuroscience. 2009; 29(6):1784–1795. doi: 10.1523/JNEUROSCI.4452-08.2009 PMID: 19211885

36. Hess EH, Polt JM. Pupil size in relation to mental activity during simple problem-solving. Science. 1964; 143(3611):1190–1192. doi: 10.1126/science.143.3611.1190 PMID: 17833905

37. Bradshaw J. Pupil size as a measure of arousal during information processing. 1967.

38. Payne DT, Parry ME, Harasymiw SJ. Percentage of pupillary dilation as a measure of item difficulty. Perception & Psychophysics. 1968; 4(3):139–143. doi: 10.3758/BF03210453

39. Arrighi R, Lunardi R, Burr D. Vision and audition do not share attentional resources in sustained tasks. Frontiers in Psychology. 2011; 2. doi: 10.3389/fpsyg.2011.00056 PMID: 21734893

40. Wahn B, König P. Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. Frontiers in Psychology. 2015; 6. doi: 10.3389/fpsyg.2015.01084 PMID: 26284008

41. Cavanagh P, Alvarez GA. Tracking multiple targets with multifocal attention. Trends in Cognitive Sciences. 2005; 9(7):349–354. doi: 10.1016/j.tics.2005.05.009 PMID: 15953754

42. Pylyshyn ZW, Storm RW. Tracking multiple independent targets: Evidence for a parallel tracking mechanism. Spatial Vision. 1988; 3(3):179–197. doi: 10.1163/156856888X00122 PMID: 3153671

43. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4; 2014.

44. Twisk JWR. Applied Multilevel Analysis. Cambridge University Press; 2006.

45. Laeng B, Endestad T. Bright illusions reduce the eye's pupil. Proceedings of the National Academy of Sciences. 2012; 109(6):2162–2167. doi: 10.1073/pnas.1118298109

46. Ahern S, Beatty J. Pupillary responses during information processing vary with Scholastic Aptitude Test scores. Science. 1979; 205(4412):1289–1292. doi: 10.1126/science.472746 PMID: 472746

47. Ahern S, Beatty J. Physiological evidence that demand for processing capacity varies with intelligence. In: Intelligence and learning. Springer; 1981. p. 121–128.

48. Wahn B, Schwandt J, Krüger M, Crafa D, Nunnendorf V, König P. Multisensory teamwork: Using a tactile or an auditory display to exchange gaze information improves performance in joint visual search. Ergonomics. 2015. PMID: 26587687

49. Wahn B, König P. Attentional resource allocation in visuotactile processing depends on the task, but optimal visuotactile integration does not depend on attentional resources. Frontiers in integrative neuroscience. 2016; 10. doi: 10.3389/fnint.2016.00013 PMID: 27013994

50. Wahn B, Schmitz L, König P, Knoblich G. Benefiting from Being Alike: Interindividual Skill Differences Predict Collective Benefit in Joint Object Control. In: Proceedings of the 38th Annual Conference of the Cognitive Science Society; 2016. p. 2747–2752.

51. Tsai YF, Viirre E, Strychacz C, Chase B, Jung TP. Task performance and eye activity: predicting behavior relating to cognitive workload. Aviation, space, and environmental medicine. 2007; 78(Supplement 1):B176–B185. PMID: 17547318

52. Konstantopoulos P, Chapman P, Crundall D. Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. Accident Analysis & Prevention. 2010; 42(3):827–834. doi: 10.1016/j.aap.2009.09.022

53. Privitera CM, Renninger LW, Carney T, Klein S, Aguilar M. Pupil dilation during visual target detection. Journal of Vision. 2010; 10(10):3. doi: 10.1167/10.10.3 PMID: 20884468

54. Unsworth N, Robison MK. Pupillary correlates of lapses of sustained attention. Cognitive, Affective, & Behavioral Neuroscience. 2016; p. 1–15.

55. Attar N, Schneps MH, Pomplun M. Working memory load predicts visual search efficiency: Evidence from a novel pupillary response paradigm. Memory & Cognition. 2016; p. 1–12.

56. Kietzmann TC, Geuter S, König P. Overt visual attention as a causal factor of perceptual awareness. PLoS ONE. 2011; 6(7):e22614. doi: 10.1371/journal.pone.0022614 PMID: 21799920