Reading Working Memory through Pupillometry

Meinris van den Bos

Master Thesis – Cognitive Psychology and Physiology

S1576917
July 2022
Department of Psychology
University of Groningen
Examiner/Daily supervisor: Sebastiaan Mathôt

Abstract

Besides responding to light, the pupil responds to a wide variety of cognitive processes such as emotions, attention, effort, imagination. More recently it was also found that changes in pupil size reflect the contents of information that is encoded and maintained in visual working memory (VWM). However, it was found that the effects of maintaining bright vs dark information on pupil size attenuate within seconds. Recent studies have suggested that, under certain circumstances, maintenance in VWM is done in activity-silent hidden states, meaning that information is not actively maintained in neural activity patterns. As hidden states cannot be measured through traditional recording methods, a measuring paradigm called pinging was devised: by briefly flashing a task neutral visual impulse, or "ping", task-evoked neural activity is elicited, and from this activity the content of VWM can be decoded. The goal of this study is to investigate whether the contents of VWM can be revealed in the pupil through pinging. In our experiment participants memorised bright or dark stimuli. Which stimulus had to be remembered, and thus maintained in VWM, was either cued before stimuli presentation or after stimulus presentation. To reveal the contents of VWM, after memorization, a task-neutral impulse stimulus was briefly displayed. Although we did find that pupil size reflects encoding of information in VWM, we did not find significant effects for maintenance, nor did we find significant effects after pinging with an impulse stimulus. As items maintained in VWM are not reliably reflected in changes of pupil size, we can conclude that for now the hidden states of VWM remain hidden from pupillometry.

Introduction

The pupillary light reflex

One of the simplest visually evoked responses is the constriction of the pupil through light (Loewenfeld, 1993). Looking at a bright object causes the constriction of the pupil (i.e. the diameter of the pupil becomes smaller). A dark stimulus, on the contrary, leads to the dilation of the pupil (i.e. the diameter of the pupil becomes larger). This phenomenon is referred to as the pupillary light response (PLR), also known as the pupil light reflex. The PLR, mediated by the autonomic nervous system (ANS), is thought to be a mechanism that helps vision adapt to the levels of available light (Campbell & Gregory, 1960). However, what is the precise role of the PLR in aiding vision?

The role of pupil dilation is quite obvious. Namely, a dilated pupil allows for more light to shine on the retina, which allows for more information being processed. This is especially relevant in the case of low-light scenarios, where a dilated pupil is essential to adequately capture contrasts and faint stimuli (Mathôt & Van der Stigchel, 2015). On the other hand, the role of pupil constriction is less evident. Several advantages have been postulated. Firstly, constriction of the pupil reduces the surface of the lens (together with any eventual lens aberrations), which increases visual acuity by reducing peripheral blur (Campbell & Gregory, 1960; Liang *et al.*, 1997). Secondly, a constricted pupil is believed to aid in transitioning from light to dark. More specifically, exposure to light causes the rods and cones to become bleached, or light adapted, rendering them insensitive. As dark adaptation takes a long time, exposing the cones to less light by constricting the pupil can help with a swifter transition from light to dark whenever necessary.

Interestingly, the pupil falls under the influence of more factors than merely that of light. For instance, even when the intensity of light remains constant, fluctuations in the diameter of the pupil may be observed as a result of the autonomic nervous system (ANS) (Loewenfeld, 1993). The ANS is a system that regulates largely subconscious physiological processes, such as - but not limited to - heart rate, digestion, arousal, and, relevant to this thesis, pupillary response. The ANS can be divided into a sympathetic and a parasympathetic system. The sympathetic system is associated with wakefulness, arousal, stress, and the fight and flight response and causes pupil dilation mediated through the pupil dilator muscle. The parasympathetic system, which is conversely associated with ”rest and digest”, causes pupil constriction mediated by the iris sphincter muscle. The sympathetic and parasympathetic systems are controlled both directly and indirectly by higher systems, i.e. hypothalamus, brainstem nuclei, the amygdala, the nucleus solitarius, as well as numerous regions of the limbic cortex. As such, the complex interplay between the sympathetic and parasympathetic activity leading to either constriction or dilation of the pupil, may reflect aspects of the cognitive and emotional state of the individual (Binda & Murray, 2015; Mathôt, 2018)

Another influence on pupil size is cognitive load. Cognitive load refers to the amount of working memory resources used (Sweller, 2011). For example, Kahneman & Beatty (1966) have shown that the pupil dilates when items are being kept in working memory. In their study participants were asked to memorise a variable number of digits. They then observed that pupil size reflected the number of remembered digits; the harder the task (more digits), the larger the pupil.

Additionally, recent evidence suggests that the PLR also appears to be modulated by higher cognitive processes such as attention (Binda *et al*., 2013a), mental effort: how *hard* we pay attention (Beatty, 1982; Karatekin, 2004), subjective interpretation (Laeng & Endestad, 2012; Naber *et al*., 2013), and mental imagery (Binda *et al*., 2013b; Laeng & Sulutvedt, 2014). Binda *et al*. (2013a) showed that covertly shifting attention towards bright surfaces constricted pupil size based on the brightness of said image. Covert shifting of attention is focusing attention to specific regions of image without moving the eyes. To measure this, they asked participants to shift their attention to brighter regions of an image, while cognitive load and retinal illumination remained constant. Through their experiments they proved that pupil constriction, but not dilation, is modulated by attention.

In a different study, the effect of contextual information from images was shown to influence pupil size. Binda *et al*. (2013b) asked participants to focus their gaze at either pictures of the sun, or control images. Even though the luminosity of both types of pictures was controlled for, the contextual information of the sun was proven to have an effect on pupil constriction beyond that of what the brightness of the pictures account for.

To further determine whether mental imagery has an effect on pupil size, Laeng & Sulutvedt (2014) asked participants to remember a triangle that could have different luminances. Subsequently participants were instructed to imagine the remembered triangle, while looking at a grey screen. Results showed that the pupil constricted or dilated in respect to the imagined bright or dark stimuli. A follow-up experiment confirmed these findings. Namely, participants were asked to look at a grey screen and imagine familiar scenarios, such as a “sunny sky” or a “dark room”. Here, they found that participants’ pupils also constricted when imagining bright scenarios compared to dark scenarios, independent of scenario complexity.

Visual working memory and its relation to the pupillary light reflex

Interestingly, mental imagery seems to be highly related to visual working memory (VWM) (Baddeley & Andrade, 2000; Keogh & Pearson, 2011; Pearson *et al*.., 2015). However, it is important to make a difference between mental imagery and VWM. Where VWM retains a memory of an item that was just seen in a state and makes it available for cognitive processing, mental imagery refers to the representation of an item without any external stimulus. Such imagery items can be recalled from long term memory, or they can be a novel combination of several recalled features (Pearson *et al*., 2015).

VWM is a cognitive system that maintains and stores a limited amount of information in mind for brief periods of time, having the information ready for immediate use. VWM allows the link between perception and long term memory (Baddeley, 2003), thus facilitating the mental manipulation of information. VWM consists of encoding, and maintenance (Baddeley, 1992).

In VWM, encoding occurs when an object that is attended to is stored in memory. Encoding has often been studied by using a pre-cue paradigm. A pre-cue *cues* the relevant properties of the upcoming task (Wang *et al*., 2017). For example, a spatial pre-cue could indicate through an arrow which stimuli needs to be remembered. Maintenance in VWM occurs when the object is no longer visibly attended, but still present (or not forgotten) in VWM. Maintenance has often been studied by use of a retro-cue paradigm. In the retro-cue paradigm, first several targets need to be memorised and maintained in working memory. During retention of the targets, a retro-cue then indicates which task relevant target stimulus is relevant (Landman *et al*., 2003; Souza & Oberauer, 2016; Zokaei *et al*., 2014).

To explore the relation between pupil size and encoding in VWM, Blom *et al*., (2016) tasked participants with memorising the shape, orientation, or exact brightness level of covertly attended stimuli. Covert attention was achieved by having participants constantly look at a fixation spot. They found that the encoding of a bright stimulus in VWM, compared to encoding of a dark stimulus leads to a pupil constriction.

How maintaining dark and bright items in VWM affects pupil size was further investigated by Hustá *et al*. (2019). They designed a paradigm to distinguish between VWM maintenance and VWM encoding. In this paradigm, participants had to memorise dark and bright stimuli, and were subsequently cued with a retro-cue to which memorised stimulus they had to recall. After the retro-cue, the change in pupil size was measured. Here it was observed that the pupil was more constricted when participants maintained a bright item, as compared to when a dark item was maintained. This finding is consistent with recent studies that show that such activity is periodical or temporary (Sreenivasan *et al*., 2014; Stokes, 2015; Wolff *et al*., 2017), and the result was reproduced by Zokaei *et al*. (2019).

When the PLR is discussed in the context of visual working memory (VWM), it is very important to make a distinction between cognitive load and the contents of VWM. In the study of Kahneman & Beatty (1966), pupil size is a reflection of cognitive load. However, as the imagery and the encoding/maintenance studies show, pupil size can also be an indication of *what* someone is thinking about.

Decoding Visual Working Memory

Until recently it was thought that working memory was maintained by actively rehearsing information, or active maintenance, represented by sustained neural activity (Baddeley, 2003). However, recent studies have shown that sustained neural activity is not required for maintaining information in working memory (Stokes, 2015).

For example, the relatively silent moments on EEG recordings between encoding and response preparation suggest that continuous rehearsal or maintenance of the content of working memory is not always required (Barak *et al*., 2010). Furthermore, Lewis-Peacock *et al*. (2012) found that activity patterns during maintenance delays only correspond to attended items in VWM. Unattended items do not show these activity patterns, even though they are still being maintained in VWM. A possible explanation is that the sustained activity often observed in working memory tasks could be due to attention (Lundqvist *et al*., 2016; Wolff*et al*., 2017).

The mechanisms of maintenance of information in working memory can be explained through activity-silent neural states, also known as hidden states (Stokes, 2015; Wolff *et al*., 2015). Hidden states are neurophysiological parameters that determine the state of the neural system, such as short term synaptic plasticity, membrane potentials, and neurotransmitter concentrations (Fujisawa *et al*., 2008; Stokes, 2015). According to this view, information is maintained in VWM through a pattern of synaptic weights, and not as an unbroken chain of neural activity. These synaptic weights are introduced after a pattern of neurons has been activated. Since the synaptic weights remain, they allow the pattern of neurons to be re-traced after a delay of activity. This synaptic trace can potentially be explained through calcium kinetics (Mongillo *et al*., 2008). According to calcium kinetics, residual calcium remains in the synapses after activity. This residual calcium then leaves a short term synaptic trace of around 2 seconds, which can in turn serve as an activity-silent short term storage buffer of specific information.

What makes these states hidden is that they are not directly observable by conventional recording methods (Stokes, 2015), as conventional recording methods measure neural activity while hidden states are activity silent. Despite being named “Hidden States”, it is still possible to infer the input-output behaviour of these activity-silent states using techniques such as TMS (Rose *et al*., 2016), or different pinging methods (Stokes, 2015).

Wolff *et al*. (2015), showed that a method called “pinging the brain” can reveal hidden states when measuring with EEG. Pinging the brain can be seen analogously to echolocation such as sonar. In sonar, sound signals, also known as pings, are sent in pulses. The reflected echoes of these sound pulses are then interpreted to map the unseen topography; When pinging the brain, the brain is “pinged” with the addition of a neutral task-irrelevant impulse stimulus. As the input pattern is held constant (always the same impulse stimulus), the differences in the output pattern can be attributed to the underlying differences in hidden states of working memory.

To see whether the contents of visual working memory can be predicted, Wolff *et al*. (2017) conducted a study where participants had to memorise randomly oriented grated stimuli. While maintaining the stimuli in VWM, the brain was retro-pinged with a task-irrelevant high contrast visual stimulus, or an “impulse stimulus” while being measured by EEG. The results show that pinging with this impulse stimulus would drive a VWM-specific impulse response, which could be measured non-invasively. The specific items being maintained in VWM were decoded from the impulse response, and it was found that these faithfully reflected both attended and unattended items. Interestingly, recently forgotten information left no traces.

Short term changes in hidden states could play an important role in high-level cognition. As the pupil size can be modulated by the contents of working memory, (Hustá *et al*., 2019; Zokaei *et al*., 2019) and top down effects from higher cognition, we expect to be able to measure the contents of hidden states of VWM through pupil size.

For our task we will also employ the idea of an "impulse stimulus" to drive neural activity in the visual system. Then we will measure the change in pupil size with pupillometry, to see whether the contents of hidden states can be read from the pupillary response.

On this basis, the scope of this thesis will be to 1) reproduce the findings of Hustá *et al*. (2019), and to 2) explore the hypothesis that pupil size reflects the information present in VWM, when retro-pinged through an impulse stimulus. Namely, after the impulse stimulus is cued, a significant difference in pupil size is expected to occur between retrieved bright and dark stimuli.

Methods

Participants

Twenty-three healthy adult students of the University of Groningen were included in the analysis of the experiment (ten female). Participation was rewarded with course credits. Participants had normal (uncorrected) vision, and could not wear glasses or lenses during the experiment. Participants were also asked to not wear makeup. The study was approved by the local ethics review board of the Department of Psychology of the University of Groningen (PSY-2021-S-0106).

Apparatus

Eye movements and the pupil size of the dominant eye were recorded with an EyeLink 1000 eye tracker, and the data was sampled at 1000hz. The experiment took place in a dimly lit room where the participant was seated and rested their head in a chin rest placed in front of a computer screen. Viewing distance was set at approximately 60cm. Stimuli were presented on a 27-inch Iiyama ProLite G2773HS computer screen with a 100hz refresh rate and a resolution of 1920 x 1080 pixels. A standard qwerty keyboard was used to log the response input from all participants.

Stimuli

Stimuli were presented as circles on a grey background (62 cd/m^2). The stimuli were always bright and dark circles that varied in luminosity, with a size of 34.6 px/°. A dark and bright stimulus were always presented together on the screen. Luminosity was randomly generated from a range between 11 cd/m^2 and 19 cd/m^2 for the dark luminosity condition and between 88 cd/m^2 and 96cd/m^2 for the light luminosity.

The brightness of the stimulus was adapted based on the accuracy of the participant's responses to change the difficulty level of the task. The brightness was maintained at 75% accuracy for dark and light stimuli separately by a three-up-one-down staircase (1d3u) procedure. The 1d3u procedure would increase the difficulty of the task after three correct responses are given in a row, and decrease the difficulty after one incorrect response has been given. The reason for the 1d3u procedure is to keep the difficulty of the task similar between the conditions for each participant, to eliminate task difficulty as a confounder on pupil size between the conditions.

Procedure

At the start of the experiment the eye tracker was calibrated using a nine-point calibration procedure. Afterwards, the participants were instructed to keep their eyes focused on the black fixation dot in the centre during each trial. There were 10 practice trials, followed by 16 blocks of 16 trials. The conditions for each trial were fully randomised. Furthermore, the participants could take a break after each block.

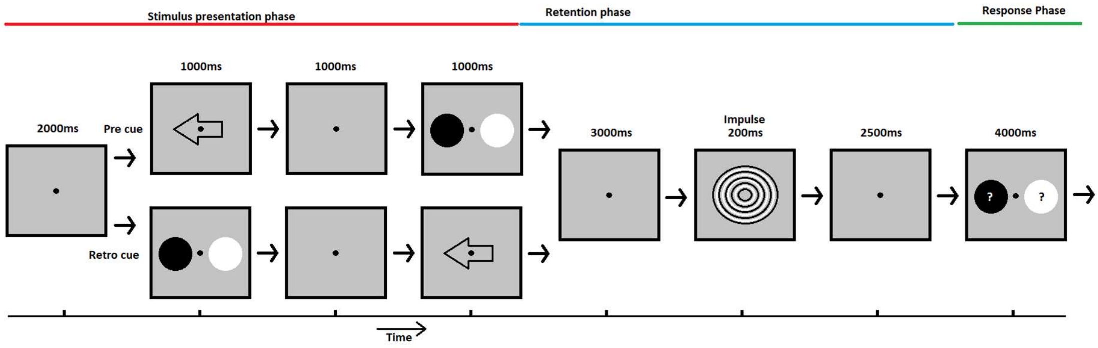


Figure 1. This figure shows the structure of the trials. Each trial starts with the stimulus presentation phase. First fixation dot is shown for 2000ms; for the pre-cue condition followed by: a cue for 1000ms, a fixation dot for 1000ms, and a stimulus screen for 1000ms; for the retro-cue condition followed by: a stimulus screen for 1000ms, a fixation dot for 1000ms, and a cue for 1000ms. The stimulus presentation phase is followed by the retention phase, which consists of a fixation dot for 3000ms, a 200ms impulse screen, followed by a 2500ms fixation dot. Finally the trial has a response phase where the participant can enter a keyboard response.

For each trial, the participant was instructed to memorise the particular brightness level of the stimuli, dark or bright circles, that appear on the grey background. The trial structure of the pre-cue and retro-cue conditions differ slightly. In both conditions a cue arrow is presented, but in the Pre-cue the arrow is presented before the stimulus, while in the retro-cue the arrow is presented after the stimulus. See Figure 1 for the structure of a trial.

In the pre-cue condition, a fixation dot was initially presented for 2000ms, followed by a cue arrow for 1000ms, followed by a 1000ms fixation dot, followed by a presentation of two stimuli for 1000ms, followed by a 3000ms fixation dot. After this, the impulse was flashed for 200ms, followed by another 2000ms fixation dot. Finally in the response-phase, a circle of a similar or different brightness than the remembered stimuli was presented. Participants had to indicate whether the circle had the same brightness by pressing the 'A' key on the keyboard, or a different brightness by pressing the 'L' Key on the keyboard.

The retro-cue condition was almost identical to the pre-cue condition, except for the first part: In the retro-cue condition after the initial fixation dot of 2000ms, the two stimuli were presented for 1000ms, followed by a 1000ms fixation dot, followed by the a cue arrow for 1000ms.

For both the pre-cue and the retro-cue condition, the participant was instructed to memorise the particular brightness level of the dark or bright circles that appeared on the grey background. In both conditions an arrow indicated which of the two stimuli would have to be remembered. In the pre-cue condition the arrow appeared before the two stimuli were presented, and in the retro-cue condition the arrow appeared after the stimuli were presented.

Results

Data pre-processing

Twenty-three participants completed a total of 5888 trials, of which 2944 represented the pre cue condition and 2944 the retro cue condition. Pupil data was corrected for blinks with the 'advanced' algorithm that is implemented in Python DataMatrix. The advanced algorithm identified blinks based on a velocity threshold, and marked data points as missing when the velocity threshold was exceeded. Then it attempted to reconstruct the missing data by interpolating the onset (start) and offset (end) of the blink. If the blink interval lasted too long (>500ms), the interval was not reconstructed and marked as missing data. In three trials, pupil size during the baseline window could not be determined due to excessive blinking; these trials were excluded from the analysis.

Next, the pupil data was down-sampled from 1000Hz to 100Hz. The pupil data was then baseline-corrected by subtracting the mean pupil size that was calculated between 3000ms and 4000ms, i.e. just before the onset of the retro cue (in the retro-cue condition) or memory stimuli (in the pre-cue condition). Subsequently, trials where the Z-score of the baseline deviated more than two standard deviations compared to the participants average baseline were excluded. This resulted in a further exclusion of 4.398% of the total trials.

Finally, for the Pre-cue condition, trials were excluded when the participants' horizontal gaze deviated too far from the central fixation point during the encoding phase, which happens on stimulus presentation, so between 4000ms and 5000ms. What would be too far from the central point was decided by calculating where the edges of the stimuli would be. According to the Eyelink 1000 Use Manual, the default Eyelink coordinates are reported in 1024 x 768 units with the coordinates (0,0) being in the top-left. The centre of the fixation dot in our study was always at the coordinates (512,384). The edges of the presented stimuli would be at least 128 units away on the x-axis from the centre of the fixation dots, and these were used as our boundaries: if the participants gaze for the trial was below x=384 units, or above x=640 units, the trial was excluded from the Pre-cue condition. After correcting for gaze, a total amount of 4948 trials out of 5888 trials (84%) remained.

Excluding participants when their gaze exceeds certain boundaries in the pre-cue condition is to prevent pupil size to be systematically biassed by looking straight at the target stimulus. The reason why gaze was not an exclusion factor for the retro-cue condition was because participants had no knowledge of which stimulus would have to be remembered during stimulus presentation, and therefore could not be systematically biassed towards the target stimulus.

Of the 4948 selected trials, for the pre-cue condition, mean accuracy was 74% for the bright trials and 73% for the dark trials. For the retro-cue condition, mean accuracy was 73% for the bright trials and 68% for the dark trials. According to the three-down one-up (3D1U) procedure, the accuracy is expected to be around 75%. Although the staircase procedure did not perform perfect in this case, with dark trials for the retro-cue having relatively many errors, it was still deemed unlikely that effects on pupil size are influenced by accuracy effects of the correct/incorrect responses in the trials.

Lastly, visual comparison of the pupil size effects between the correct vs. incorrect trials did not indicate any qualitative differences. Therefore all 4948 correct and incorrect remaining trials were included in the linear mixed effects analysis.

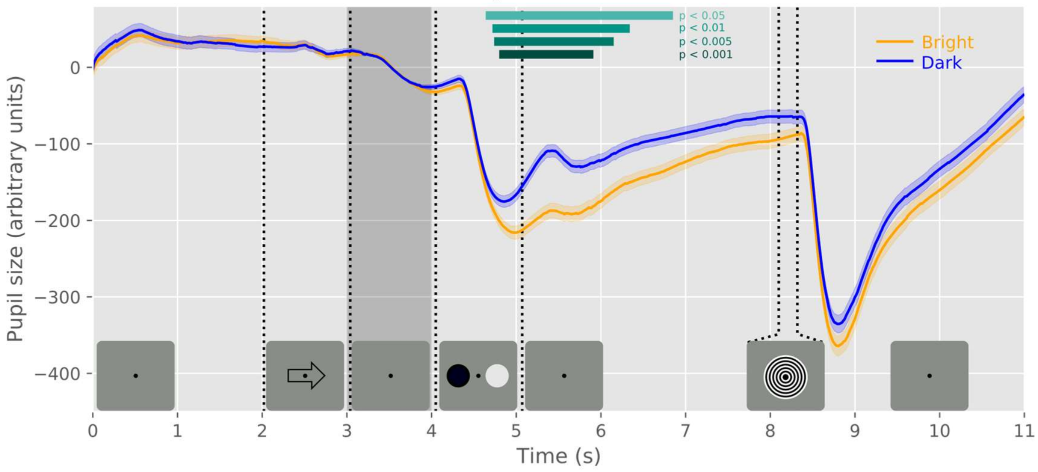
Pupillary responses – mixed effects

Next, a linear mixed effects analysis was conducted, using the R package lme4 (Bates *et al.*, 2015) on each 10ms window, with pupil size as a dependent measure, and brightness (bright/dark), and condition (pre cue/retro cue) as fixed effects. The difference in pupil size between the conditions was deemed significant when $p < 0.05$ was sustained beyond 200 consecutive milliseconds (20 windows). The main results of the experiment are shown in figure 2. The pupil sizes were plotted over time for the entire length of the trial, individually for the pre cue (a) and retro cue (b) conditions. For each condition, the plots differentiate between the bright (orange), and the dark (blue) stimulus presentation.

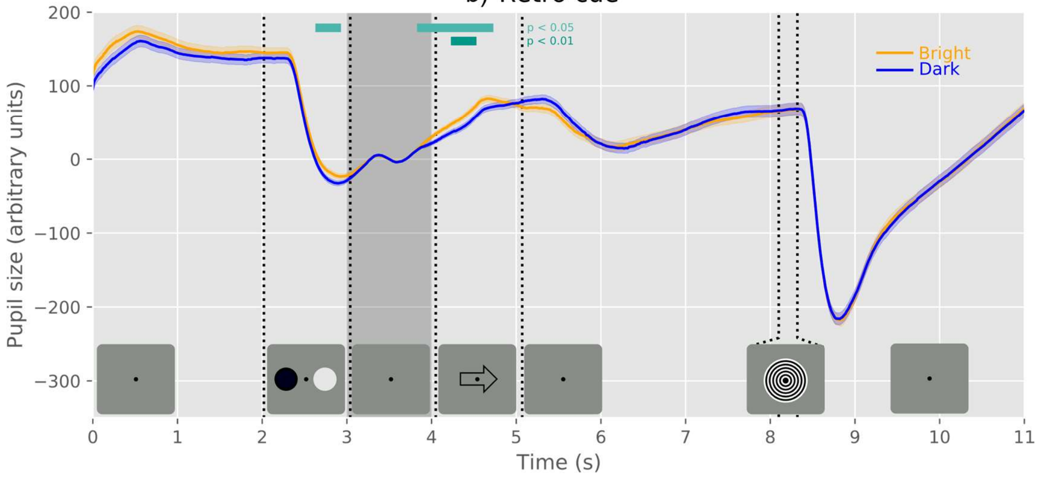
In the pre cue condition, a significant effect on pupil size was noted between 4700ms and 6790ms. This effect started during the encoding phase and persisted throughout the retention interval. This observation is in line with the results from Husta *et al*. (2019), who also noted an effect of brightness on pupil size, while participants covertly attended to the target stimuli. This observation was attributed to the encoding of the brightness of the stimulus in visual working memory.

In the retro cue condition, a significant effect was present between 3900 and 4690ms. Given the onset of this effect preceded the actual presentation of the cue, the change in pupil size is likely a spurious effect. Therefore it cannot be concluded that the pupillary response reflects content of visual working memory during maintenance.

Furthermore, after the impulse stimulus was flashed, no significant effects occurred in either the pre cue condition, or the retro cue condition.



a) Pre cue



b) Retro cue

***Figure 2*. Average pupil size over time.** Average pupil size for all participants plotted through the entire trial, for the (a) pre cue condition and the (b) retro cue condition separately. The orange lines indicate the trials for which the bright stimulus had to be remembered, and the dark blue line indicates the trials for which the dark stimulus had to be remembered. The trial structure is displayed at the bottom of each graph. The turquoise lines show where in the trial a significant effect ($p < .05$) occurred for at least 200ms consecutively. (a) For the pre cue a significant effect occurred between 4700ms and 6790ms, which can be explained by encoding effects of the brightness information. (b) For the retro cue the significant effect starts before the target is cued. This means that either participants used a memorization strategy, or the significant effects are purely based on chance.

Pupillary responses - Individual effects

To characterise the individual effects, the difference in average pupil size between the bright and dark condition was calculated for each individual, within the 4500ms and 6500ms window, for both the pre- and retro cue condition (Figure 3). In the pre cue condition, all participants except for one have shown an effect in the expected direction; namely, that memorising a bright stimulus results in, on average, a smaller pupil compared to memorising a dark stimulus. It should be noted, however, that in the case of four participants no trials remained in the pre-cue condition for either dark or bright stimuli after filtering for eye-movements and blinks. As such, the difference in averages for these participants could not be included.

For the retro-cue condition, the individual effects did not occur in a specific direction (Figure 3). Namely, for half the participants the pupil was smaller during encoding of the dark stimuli than that of the bright stimuli. This is in contrast with previous findings, where pupil size would be relatively smaller during encoding of bright stimuli.

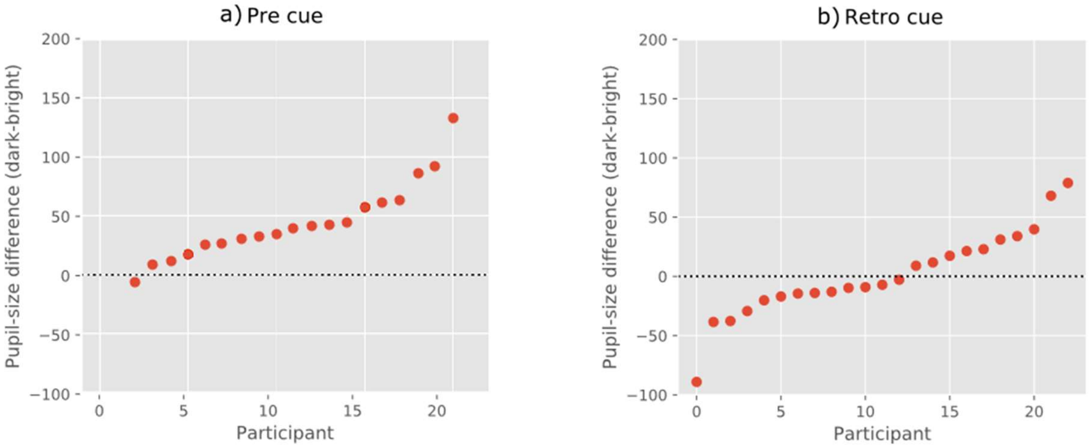


***Figure 3*. Average effects of pupillary responses for individual participants.** Average effects of the difference in mean pupil (between 4500ms and 6500ms size between the light condition and the dark condition for individual participants in the pre cue condition and the retro cue condition. The points are rank-ordered based on effect size within their condition. In the pre cue condition the effects for four participants are missing due to trials being excluded on the basis of gaze and a divergent baseline (Z > 2).

Discussion

In this study we use pupillometry to investigate whether the contents of visual working memory (VWM) are reflected in the pupillary light reflex (PLR), and whether the contents of VWM can be predicted through changes in pupil size after an impulse stimulus is “pinged”. We adapt the paradigm used by Hustá *et al*. (2019), who found that pupil size reflected the encoding and maintenance of pre-cued and retro-cued bright and dark stimuli. Namely, we extend this paradigm by adding a task-neutral impulse stimulus that is pinged for a very short time (100ms).

Here, we show that the contents of VWM are reflected in the PLR during encoding by using a pre-cue stimulus; that is, a cue that informs participants *beforehand* whether the task-relevant stimulus is bright or dark. The purpose of the pre-cue stimulus is to inform participants where to covertly focus their attention when both a bright and a dark stimulus are presented. When they covertly focus their attention on a bright or dark stimulus, the cued stimulus gets encoded in VWM. Because the bright or dark stimulus present in working memory are respectively reflected by a relatively smaller or larger pupil size, this finding shows that the PLR does reflect higher cognitive processes, specifically covert visual attention. This finding is consistent with previous studies that examined whether encoding of information in VWM is reflected in changes in pupil size (Blom *et al*., 2016; Hustá *et al*., 2019).

However, no effect for dark/ bright has been found in the retro-cue condition; that is, the condition where participants are only informed which stimulus is task-relevant *after* the stimuli are no longer visible on the display. The purpose of the retro-cue stimulus is to isolate maintenance in VWM from covert attention during encoding in VWM. Namely, when information is still present in VWM, the retro-cue signals the participant which of the two circles present in VWM have to be maintained. That we did not find an effect is contrary to earlier evidence (Hustá *et al*., 2019; Zokaei *et al*., 2019). For instance, Hustá *et al*. (2019) described a clear effect of bright and dark stimuli on changes in pupil size following the retro-cue presentation, which attenuated fairly quickly. In other words, they found that pupil size becomes significantly larger when the retro-cue indicates that the dark stimulus has to be maintained, as opposed to when the bright stimulus has to be maintained. However, we did not replicate this effect in our data.

Furthermore, we investigate whether the information being maintained in VWM can be predicted through changes in pupil size after a simple neutral impulse stimulus. When pinging, a task-neutral impulse stimulus is flashed briefly, after which the response to this impulse stimulus can be measured. Since the pinged task-neutral impulse stimulus is constant across all conditions, any effects found in the response to this stimulus should reflect the contents of VWM. In our study, no significant changes in pupil size between bright and dark stimuli are found after the impulse stimulus is pinged, which is opposed to evidence from previous electroencephalography (EEG) studies where an effect is observed (Wolff *et al*., 2015, 2017, 2020).

So why did we not observe an effect of maintaining bright and dark stimuli in VWM on pupil size, whereas Hustá *et al*. (2019) and Zokaei *et al*. (2019) did? There are several possible explanations. First, one reason could be that participants simply forgot the stimuli, or failed to do the task. However, because participants performed around the expected accuracy of +-75% in the retro-cue condition, we can assume that the cued stimuli were being maintained in VWM until the end of the trials.

Second we wonder if it is possible that participants were maintaining both the dark and the bright stimulus, regardless of the instruction to forget the irrelevant stimulus. In this case, the combined effects of maintaining a dark and bright stimulus would cancel each other out. Participants were explicitly instructed and reminded that they only have to attend to the stimulus that is cued, and that they can forget the stimulus that is no longer task-relevant. Wolff *et al*. (2017) have shown that VWM is highly flexible, and representations of specific items can be rapidly cleared when they are no longer relevant to the task through directed forgetting. Therefore we assume that participants did correctly attend and maintain the cued stimulus, and automatically “forgot” the unattended stimulus as it was no longer task-relevant.

A third explanation for the lack of a dark/bright effect during maintenance is related to the concept of hidden states, or activity-silent states (Stokes, 2015). Hidden states are named as such, simply because they cannot directly be measured by conventional recording methods. The mechanism of hidden states can be explained by the short-term synaptic plasticity of the neurons involved in VWM (Barak *et al*., 2010; Zucker & Regehr, 2003). According to this view, information is maintained as a pattern of synaptic weights, instead of as an unbroken chain of neural activity. After a pattern of neurons has been activated, a short-term synaptic trace is left behind which can be reactivated. A possible mechanism that explains synaptic weights is calcium kinetics (Mongillo *et al*., 2008), where residual calcium in synapses could leave a synaptic trace and serve as a storage buffer of specific information. This would allow information to remain in VWM as a hidden state (Stokes, 2015). As the maintenance of the luminosity of stimuli might also be a hidden state, the attended maintained stimulus might simply not have engaged the pathways that modulate pupil size frequently enough for us to have found an effect of brightness. Due to the nature of hidden states, maintenance of the attended stimulus does not require continuous neural activity, and consequently is not reflected in changes in pupil size. However, an important consideration is a possibility that for Hustá *et al*. (2019) and Zokaei *et al*. (2019) the pathways that modulate pupil size *were* in fact engaged frequently enough while maintaining stimuli in VWM. Therefore, the possibility that hidden states are reflected in pupil size cannot be disregarded.

Fourth, although unlikely, it is still possible that the effects of brightness on pupil size during maintenance that were found by Hustá *et al*. (2019) and Zokaei *et al*. (2019) are effects of conscious re-representations during maintenance. We deem this unlikely as both the task and corresponding instructions of Hustá *et al*. (2019) are very similar to ours. Nevertheless, it is possible that the attended maintained item was consciously rehearsed in VWM during the presentation of the retro-cue. When the stimulus is no longer actively rehearsed, and no mental imagery takes place as a strategy to consciously “keep” the item active in VWM, it is possible that no effects of maintenance could be observed in changes in pupil size. Imagery is a conscious re-representation of a perception, and imagery has been shown to adjust pupil size. Imagining a bright object leads to pupil constriction, as opposed to imagining a dark object (Laeng & Sulutvedt, 2014). Neuroimaging studies have shown that during imagery of an object, the pattern of activity within the visual cortex is nearly identical to the pattern during the perception of said object. Mental imagery can be used for mnemonic performance (Pearson *et al*., 2015) by visualizing stimuli and thus keeping them active in VWM. Hustá *et al*. (2019) report that participants employed both visualization and verbalization strategies for remembering brightness information. In our study, we did not ask participants about which strategies they applied, so it is possible that participants employed a mix of verbalization and visualization strategies, meaning they sometimes could have employed mental imagery, and sometimes verbalization strategies. If participants in our study did not consciously keep up visualization through mental imagery, or if participants switched between strategies, it is possible that changes in pupil size could not significantly reflect the maintained contents of VWM.

Lastly, another explanation for not replicating the results of Hustá *et al*. (2019) might also be that the effect of attending the dark and bright stimuli on pupil size during maintenance is not strong enough to be detected. It is possible that the effects of brightness on pupil size during maintenance are being drowned out by other noise, and therefore did not reach significance in our study.

No significant effect for pupil size was found between the bright and dark condition after pinging an impulse stimulus. The purpose of the impulse stimulus is to reveal the contents of VWM that are being maintained within an activity-silent state, also known as a *hidden state*. As we did not find an effect in pupil size during the maintenance of dark and bright stimuli, we can assume that pinging the maintained stimuli in VWM would also not cause a bright/dark effect in pupil size. As such, it can be concluded that changes in pupil size following pinging with an impulse stimulus likely do not reveal the contents of hidden states in VWM.

Thus, could activity-silent states influence the pupil? According to Christophel *et al*. (2017), a diverse set of brain regions are involved in maintenance in VWM. This, amongst others, includes the prefrontal (Lara & Wallis, 2015), visual and parietal cortex, as well as the medial temporal lobe (Kamiński *et al*., 2017), and subcortical regions such as the superior colliculus and the thalamus. The exact location of the hidden states in the brain are difficult to pinpoint due to the distributed nature of the system. The *distributed systems view of working memory* perspective (Lorenc & Sreenivasan, 2021) views the functions of VWM as distributed in parallel, across multiple brain regions, which are then modulated by context and task goals. In this perspective, it is possible that the hidden states of VWM in our task do not, or barely, include the areas of the brain that can modulate the dilation or constriction of the pupil based on the brightness of the attended maintained stimulus.

To conclude, although no effect on pupil size was found during maintenance and after pinging, a clear effect of brightness on pupil size was found for encoding of items in VWM. This shows that the PLR can reflect higher cognition and encoding in VWM. However, it should be borne in mind that the higher cognitive functions having an influence on pupil adaptation, may not, or barely be involved in the network that maintains brightness information in VWM through activity-silent states. More insight into the processes underlying the effect of cognition on pupil size is needed to better understand both high level cognition and the processes of VWM. Furthermore, whether the pupil is a suitable vessel to read out the contents of visual working memory through a probe remains to be seen.

References

Baddeley, A. (1992). Working Memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/SCIENCE.1736359

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Barak, O., Tsodyks, M., & Romo, R. (2010). Neuronal population coding of parametric working memory. *Journal of Neuroscience*, *30*(28), 9424–9430. https://doi.org/10.1523/JNEUROSCI.1875-10.2010

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

Binda, P., & Murray, S. O. (2015). Keeping a large-pupilled eye on high-level visual processing. *Trends in Cognitive Sciences*, *19*(1), 1–3. https://doi.org/10.1016/j.tics.2014.11.002

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

Binda, P., Pereverzeva, M., & Murray, S. O. (2013b). Pupil constrictions to photographs of the sun. *Journal of Vision*, *13*(6), 1–9. https://doi.org/10.1167/13.6.8

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*(11), 1716–1723. https://doi.org/10.1037/xhp0000252

Campbell, F. W., & Gregory, A. H. (1960). Effect of Size of Pupil on Visual Acuity. *Nature*, *187*(4743), 1121–1123. https://doi.org/10.1038/1871121c0

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J. D. (2017). The Distributed Nature of Working Memory. *Trends in Cognitive Sciences*, *21*(2), 111–124. https://doi.org/10.1016/j.tics.2016.12.007

Fujisawa, S., Amarasingham, A., Harrison, M. T., & Buzsáki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nature Neuroscience*, *11*(7), 823–833. https://doi.org/10.1038/nn.2134

Hustá, C., Dalmaijer, E., Belopolsky, A., & Mathôt, S. (2019). The Pupillary Light Response Reflects Visual Working Memory Content. *Journal of Experimental Psychology: Human Perception and Performance*, 1–7. https://doi.org/10.1037/xhp0000689

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science (New York, N.Y.)*, *154*(3756), 1583–1586.

Kamiński, J., Sullivan, S., Chung, J. M., Ross, I. B., Mamelak, A. N., & Rutishauser, U. (2017). Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nature Neuroscience*, *20*(4), 590–601. https://doi.org/10.1038/nn.4509

Karatekin, C. (2004). Development of attentional allocation in the dual task paradigm. *International Journal of Psychophysiology*, *52*(1), 7–21. https://doi.org/10.1016/j.ijpsycho.2003.12.002

Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE*, *6*(12). https://doi.org/10.1371/journal.pone.0029221

Laeng, B., & Endestad, T. (2012). Bright illusions reduce the eye's pupil. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(6), 2162–2167. https://doi.org/10.1073/pnas.1118298109

Laeng, B., & Sulutvedt, U. (2014). The Eye Pupil Adjusts to Imaginary Light. *Psychological Science*, *25*(1), 188–197. https://doi.org/10.1177/0956797613503556

Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, *43*(2), 149–164. https://doi.org/10.1016/S0042-6989(02)00402-9

Lara, A. H., & Wallis, J. D. (2015). The role of prefrontal cortex in working memory: A mini review. *Frontiers in Systems Neuroscience*, *9*(DEC), 1–7. https://doi.org/10.3389/fnsys.2015.00173

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*(1), 61–79. https://doi.org/10.1162/jocn_a_00140

Liang, J., Williams, D. R., & Miller, D. T. (1997). Supernormal vision and high-resolution retinal imaging through adaptive optics. *Journal of the Optical Society of America A*, *14*(11), 2884. https://doi.org/10.1364/josaa.14.002884

Loewenfeld, I. E. (1993). *The Pupil: Anatomy, Physiology, and Clinical Applications, Volume 1*. Iowa State University Press, 1993.

Lorenc, E. S., & Sreenivasan, K. K. (2021). Reframing the debate: The distributed systems view of working memory. *Visual Cognition*, *29*(7), 416–424. https://doi.org/10.1080/13506285.2021.1899091

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L. L., Buschman, T. J. J., & Miller, E. K. K. (2016). Gamma and Beta Bursts Underlie Working Memory. *Neuron*, *90*(1), 152–164. https://doi.org/10.1016/j.neuron.2016.02.028

Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, *1*(1). https://doi.org/10.5334/joc.18

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*(5), 374–378. https://doi.org/10.1177/0963721415593725

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic Theory of Working Memory. *Science*, *319*(5869), 1543–1546. https://doi.org/10.1126/science.1150769

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology*, *4*(DEC), 1–12. https://doi.org/10.3389/fpsyg.2013.00919

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, *19*(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136–1139. https://doi.org/10.1126/SCIENCE.AAH7011/SUPPL_FILE/ROSE.SM.PDF

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, and Psychophysics*, *78*(7), 1839–1860. https://doi.org/10.3758/s13414-016-1108-5

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*(2), 82–89. https://doi.org/10.1016/J.TICS.2013.12.001

Stokes, M. G. (2015). “Activity-silent” working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405. https://doi.org/10.1016/j.tics.2015.05.004

Sweller, J. (2011). Cognitive Load Theory. In *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 55). Elsevier Inc. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Wang, B., Yan, C., Wang, Z., Olivers, C. N. L., & Theeuwes, J. (2017). Adverse orienting effects on visual working memory encoding and maintenance. *Psychonomic Bulletin and Review*, *24*(4), 1261–1267. https://doi.org/10.3758/s13423-016-1205-4

Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*, *9*(september), 2–11. https://doi.org/10.3389/fnsys.2015.00123

Wolff, M. J., Jochim, J., Akyürek, E. G., Buschman, T. J., & Stokes, M. G. (2020). Drifting codes within a stable coding scheme for working memory. *PLoS Biology*, *18*(3). https://doi.org/10.1371/journal.pbio.3000625

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871. https://doi.org/10.1038/nn.4546

Zokaei, N., Board, A. G., Manohar, S. G., & Nobre, A. C. (2019). Modulation of the pupillary response by the content of visual working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(45), 22802–22810. https://doi.org/10.1073/pnas.1909959116

Zokaei, N., Heider, M., & Husain, M. (2014). Attention is required for maintenance of feature binding in visual working memory. *Quarterly Journal of Experimental Psychology*, *67*(6), 1191–1213. https://doi.org/10.1080/17470218.2013.852232

Zucker, R. S., & Regehr, W. G. (2003). Short-Term Synaptic Plasticity. *Http://Dx.Doi.Org/10.1146/Annurev.Physiol.64.092501.114547*, *64*, 355–405. https://doi.org/10.1146/ANNUREV.PHYSIOL.64.092501.114547

Supplementary Materials

Preprocessing of data

First, the naming conventions in the raw edf files had to be preprocessed using a script written in c# (due to performance reasons) (Consult Supplementary Materials), since the "start_phase retention_interval" and "end_phase retention_interval" phase tags were present twice per trial. Therefore, through an algorithm that looped through the data, for each trial the second "start_phase retention_interval" and "end_phase retention_interval" tag was renamed to "start_phase retention_interval_1" and "end_phase retention_interval".