Attention is in the eye of the beholder: An investigation of the relationship between attentional lapses and pupil size

Felea, Alexandra

Citation

Felea, A. (2021). *Attention is in the eye of the beholder: An investigation of the relationship between attentional lapses and pupil size*.

Attention is in the eye of the beholder

An investigation of the relationship between attentional lapses and pupil size

Alexandra Felea

Master Thesis Cognitive Neuroscience (Research)
Date: 22.01.2021
Student number: 2227584
Supervisor: Sander Nieuwenhuis
Second reader: Wouter Boekel

Acknowledgements

To say that the past three years of my Master have been adventurous would be a gross understatement and I cannot express my gratitude enough for being walked through this phase in my life by the best team I could have wished for.

I would like to thank my mentor and supervisor, Sander Nieuwenhuis, whose kindness and dedication have helped inspire me and restore my hope in research so many times in the past years. Thank you for all the impromptu meetings when I just needed someone to talk to about academia and for helping me finish this thesis in time for my next adventure.

I would also like to thank Verónica Mäki-Marttunen for giving me the chance to work together with her on this project and for always being there to answer all my questions with so much grace and calm even when I probably drove her crazy. Thank you, Veronica for somehow always finding the time for me even though everything must have been so new for you too when you arrived in Leiden.

I would also like to thank my family for always supporting me and making my being here possible and to thank my partner, Luka Velinov, who is my inexhaustible source of puns and programming tips and whom despite going through all the same milestones and struggles as me has always found the energy to cheer me up when I was feeling down.

Last but not least I would like to thank Evert Dekker and Iris Spruit for their invaluable help with all lab and technical related things. Thank you, Iris for your code and for taking the time to come check the equipment in the middle of a pandemic so that I could start testing and thank you, Evert for being my go-to person in these past three years for all technical problems and for putting up with me asking questions all the time and always urgently needing something.

Abstract

Attention plays a vital role in helping the brain adaptively disambiguate between relevant and irrelevant information in the environment. Lapses in attention can thus have important negative consequences, varying from small mishaps to life-threatening mistakes and as such it is important to study their mechanisms. Neural entrainment appears to play an important role in attention and researchers have long tried to explore the relationship between the two. Unfortunately, a majority of studies on the topic relies heavily on animal and clinical studies, often using invasive measurement techniques. What is more, some of the non-invasive methods used (such as eye-tracking), sometimes offer contradicting results and lack methodological consensus. The current work is part of a larger study looking at the relationship between attentional lapses and neural entrainment using a multimodal oddball task. Here, we focused solely on eye tracking and set out to investigate whether lapses in attention correspond to changes in tonic pupil size and whether measures of attentional lapses suffer from time-on-task effects. Participants performed an oddball dual-modality task in which they were presented with two simultaneous streams of stimuli (visual and auditory) and had to attend and respond to either visual or auditory targets. We found that the number of false alarms reduced significantly over time and that pupil size showed a decreasing (albeit not significant) trend, while the number of hits did not appear to decrease as time passed. These findings seem to be in partial consensus with previous research on the topic. We also found a significant difference in pupil size between the two attending modalities (attend visual and attend auditory), suggesting that visual and auditory stimuli may influence attention (or at the very least pupil diameter) differently. There was however no difference in any of the behavioral measures based on the attended modality. Similarly, no effect of the pupil size was found for our behavioral measurements, suggesting that pupil size does not necessarily predict behavior. It must be noted however that while our results appear to contradict previous findings from the literature, our sample size is likely too small to draw any generalizable conclusions from.

Key words: neural entrainment, attention, pupil diameter

Contents

Attention is in the eye of the beholder

An investigation of the relationship between attentional lapses and pupil size

The present study initially set out to explore the relationship between neural entrainment and attentional lapses using EEG and eye-tracking methods. Unfortunately, due to research delays caused by the current pandemic as well as unexpected confounds that became obvious while running the study, the data that was used in the analysis of this study will be considered a pilot for the final project. The time available for the completion of this thesis was shortened due to these circumstances. Therefore, it was decided that I would be focusing solely on eye-tracking. This change in research focus happened late in the thesis process and as such some elements of my introduction may lean more towards the entrainment topic that unfortunately will not be developed on further on.

Entrainment and Cognition

The ways in which the human brain relates to its environment are highly complex and heavily rely on the brain's ability to disambiguate between relevant and irrelevant information. While the field of cognitive neuroscience has been extensively studying the brain's mechanisms for decades, a shift away from an overemphasis on event evoked signals (where background activity is largely seen as noise) and towards a deeper exploration of self-generated neural activity has been brewing in the past years (Raichle, 2010). It is well known by now that the brain operates in a highly rhythmic manner, dominated by complex interplays of oscillations which are influenced by both endogenous and exogenous factors. These oscillations play a crucial role in extracting, interpreting and maintaining information at a perceptual level (Ohki & Takei, 2018).

Human perception is thus not compiled as a continuous stream of sensory information but rather as a series of discrete samples acquired at a constant rate. This sampling periodicity leads to a phenomenon through which rhythmic sensory inputs trigger synchronized oscillatory rhythms throughout the brain, essentially tuning it to incoming information by recreating external rhythms internally. This phenomenon is known as neural entrainment and has been defined as "*the alignment of one or more oscillating systems [here: the brain] to an external rhythm, whereby the interactions are unidirectional, that is, the external rhythm influences the oscillating system but not vice versa*" (Lakatos et al., 2019). Entrainment ensures that information is attended, perceived and understood appropriately across a large range of modalities, including auditory (Ding et al., 2017; Giraud & Poeppel, 2012) and visual (Mathewson et al., 2009; VanRullen et al., 2006). Important differences when it comes to entrainment in different modalities do exist, however, and they are modulated by the fact that different systems natively acquire information at different frequencies and at different timepoints. In the visual system for example information is typically acquired at a rate of 7 to 13 Hz, relatively uniformly across most perceptual areas while in the auditory system these oscillations appear to influence perception at specific timepoints in the acquisition timeline, particularly after feature extraction has taken place (VanRullen et al., 2014). These differences in system characteristics must be considered when comparatively exploring entrainment in different modalities.

Neural entrainment does not only shape perception, however, but also profoundly affects a wide range of other cognitive processes. Starting from childhood, humans engage in what is known as behavioral coordination, a process by which we align our behaviors and cognitive states to those of others. Behavioral coordination co-occurs with neural entrainment between individuals, which led researchers to assume that the synchronous behavior was the cause of synchronous oscillations. However, even in the absence of behavioral coordination neural entrainment still manifests in social circumstances, suggesting that shared cognitive states may also involve shared brain activation patterns (Wass et al., 2020). In adulthood, neural entrainment has been shown to be particularly important when it comes to balancing self-other integration and segregation. Periods of alpha suppression in the right centro-parietal regions have been associated with increased behavioral entrainment which promotes self-other integration. On the other hand, periods of alpha enhancement were found to correspond to self-other segregation and reduced entrainment (Novembre et al., 2016).

It becomes evident that neural entrainment and its ubiquity across cognitive domains is a driving mechanism of human cognitive functioning. It shapes the ways in which we make sense of our environment and enhances our interactions with others. What is more, it might even play a causal role in how we form and consolidate memories (Hanslmayr et al., 2019), tell time (van Wassenhove, 2016), and understand spoken language (Riecke et al., 2018).

Attention

Of all the cognitive processes that may be intrinsically linked to neural entrainment, we are specifically interested in attention. Originally, we set out to explore the relationship between neural entrainment and attentional lapses. Unfortunately, we were not able to hold that goal in the given timeframe. Our interest in attention is driven by the fact that it appears to be a somewhat understudied process in relationship to entrainment, but one that plays a key role in normal human functioning. Attention is a vital process tasked with teasing apart relevant and irrelevant information, in a brain that is constantly bombarded with both internal and external cues. While the difference in oscillatory patterns between different types of attention (e.g., sequential and parallel) remains to be further investigated (Dugué et al., 2015), consensus exists that – much like other processes – attention appears to be in general heavily modulated by neuro-oscillations in the delta and theta ranges (Landau & Fries, 2012).

The current study is based off the 2016 study of Lakatos et al. which explored the global, long-scale dynamics of neural entrainment and attention in the primary auditory cortex of macaque monkeys. In their study, monkeys were fitted with deep brain recording electrodes and were monitored while doing an intermodal selective attention task in which they had to respond to an odd-one-out type stimulus in a signaled modality (either visual or auditory). They found that periods of entrainment to the attended modality were linked to correct behavioral responses. On the other hand, attentional lapses were characterized by a lack of both behavioral responses and entrainment to either modality as well as by the occurrence of high-amplitude alpha oscillations. Moreover, periods of entrainment and lack thereof were found to be anticorrelated and to fluctuate rhythmically, with lapses of entrainment regularly appearing at 10-30 second intervals. This suggests the existence of an internal attention-resetting mechanism. Lakatos' findings are in accordance with existing literature: Alpha oscillations have been hypothesized to play an important role as a suppression mechanism in selective attention (Foxe & Snyder, 2011) by acting as a phase reset. These phase resets may prevent the brain from becoming over fixated on potentially irrelevant stimuli and facilitate information transfer between brain areas (Voloh & Womelsdorf, 2016). Phase resetting may thus fulfil a dual role in attention by maintaining a flexible neural state while at the same time facilitating the coordinated transfer of information within the brain.

A common theme that emerges from both Lakatos' and the social entrainment studies is that alpha activity suppresses entrainment and promotes flexibility whereas a lack of it maintains attention and focus. It must be noted, however, that while alpha-induced phase resetting might be a working mechanism in attentional switching and selection, it is likely not the only one. Several studies have found that overt behaviors (such as saccadic eye movements) may also trigger phase resetting (Hafed et al., 2015). Some even suggested that covert and overt attentional sampling may influence neural gain in different but interdependent ways, with overt behaviors signaling the need to re-sample from the environment and redirect behavior (for a review see Helfrich et al., 2019). Nevertheless, while overt behaviors are not the focus of this thesis, it is important to remain fully aware of the fact that other mechanisms play an important role in attentional selection too.

Considering the importance of entrainment for attending, perceiving, and understanding incoming information, it is somewhat natural to consider that a lack of neural entrainment or disorganized patterns of brain activity may lead to important deficits in these functions. Many previous studies have focused heavily on the effect of abnormal entrainment on attention and found that patients with schizophrenia show less entrainment to stimuli than healthy participants do and thus have difficulty identifying deviant tones in a sequence (Lakatos et al, 2013). Similar results were obtained in other disfunctions including, but not limited to dyslexia (Goswami, 2011) and ADHD (Lenz et al., 2010).

While the relationship between lapses in entrainment and lapses is attention in humans has been studied to some extent, primarily to inform better approaches to mental disorders, research on healthy participants remains scarce and as such the mechanisms are still somewhat unclear. Given the difficulties of confidently formulating conclusions based on animal or clinical research, our aim is to bridge this knowledge gap and to investigate the effect of lapses in entrainment on attention in healthy participants. Moreover, traditional means of measuring entrainment (deep brain recordings and EEG) may be difficult to implement in non-clinical populations and perhaps impossible to use outside of the lab due to being cumbersome or invasive. Therefore, we aim to also explore whether these relationships can be picked up using non-invasive eye-tracking techniques.

Pupil Diameter

Pupil diameter has become a staple measure of attention, with well rooted tradition in contemporary neuroscience, and for good reason. It is known to quantify cognitive load and attentional performance (Kahneman & Beatty, 1966), predict changes in exploration-exploitations paradigms (Jepma & Nieuwenhuis, 2011), and has been shown to also function as an index of attentional control state fluctuations (Gilzenrat et al., 2010). Pupil diameter is a versatile tool and a good fit for our current study as it is a reliable index of locus coeruleus-noradrenergic (LC-NA) system activity both at rest and whilst performing oddball tasks (Murphy et al., 2014). Brain areas such as the parietal cortex and the superior colliculus, which are both associated with attentional processes are heavily innervated by LC-NE offshoots (Sara & Bouret, 2012). Noradrenaline has been identified as an important neuromodulator in attentional control (A. Smith & Nutt, 1996) and it has been hypothesized that it influences attention both by interacting with arousal levels (Coull et al., 2004) and by potentially focusing attention through canceling out the effect of distractors (A. P. Smith et al., 1992).

While, as outlined above, pupil diameter has been proven to be a useful tool, some researchers have observed that studies on attention and pupil diameter offer somewhat inconsistent results in what the indicators of attention or lapses of attention are. Both findings indicating that small baseline pupil diameter is indicative of poor performance (Van Orden et al., 2000), as well as those suggesting the complete opposite (Unsworth & Robison, 2016) are quite common. Moreover, several studies have even pointed to the fact that both large and small pupil sizes are related to poor task performance (Murphy et al., 2011). Part of the explanation for such contrasting findings appears to be that studies generally use very different methodologies that might measure different aspects of attention. A more compelling explanation for these inconsistencies comes from studies such as that of Van Den Brink et al. (2016), who investigated the effect of time-on-task on pupil diameter in an attentional task. They found that depending on whether time-on task is accounted for during analysis the relationship between pupil diameter and performance behaves either linearly (with time-on-task included) or non-linearly, as an inverted U shape (when time-on task is controlled for). These findings fall in line well with the adaptive gain theory of LC-NE function (Aston-Jones & Cohen, 2005) and suggest that a Yerkes-Dodson U-curve (Yerkes & Dodson, 1908) may offer a more comprehensive account of attention, provided that researchers make sure to correct for potential confounding effects of time-on-task.

Aims and Hypotheses

In summary, neural entrainment appears to play an important role in attention, but the relationship between the two is seldomly examined using healthy human population samples and non-invasive measurement techniques. The poverty of such research means that conclusions drawn purely from animal and clinical samples may not be well generalizable. What is more, although noninvasive tools for tracking attentional lapses do exist, special care must be taken when interpreting the results due to potential confounds, such as the time-on-task effects. Having easily accessible physiological markers of attentional lapses has implications well beyond simply pinpointing the mechanisms of attention. These findings could be used to design interventions and tools to predict attentional lapses in the real world, with potentially life-saving applications.

One such application could be incorporating eye-tracking equipment in cars that would continually monitor a driver's pupil size and alert them when lapses in attention occur. Systems using pupil diameter as a measure of cognitive load during driving have already been successfully tested in simulations and show promising results for future applications (Palinko et al., 2010). It is only natural to imagine that using them for monitoring attentional states may also be possible.

We believe that it is important to explore whether the relationship between attention and neural entrainment can be detected using non-invasive techniques (namely EEG and pupil recordings). To this end, we employed a dual-modality oddball task based off Lakatos et al. (2016) in which participants were presented with simultaneous streams of visual and auditory stimuli. Participants were tasked to attend to one of the modalities at a time and respond to the rare (oddball) stimuli in that modality, but not the other.

In the following, I will focus entirely on the eye-tracking results obtained from our pilot study and will approach them in an exploratory manner. We hypothesize that:

1. Participants' mood and arousal would decrease over the course of the experiment.
2. Participants would give more false alarms and fewer hits over time.
3. Behavioral decrements (low hits, high false alarms) would correspond to changes in pupil size. Because of the contrasting results obtained by previous researchers we decided not to define the direction of this relationship but approach it in an exploratory manner.
4. There would be no difference in pupil size nor behavior between the visual and the auditory attend condition.

Due to the unforeseen occurrence of a visual illusion that potentially confounded our results we also formulated a post-hoc hypothesis:

5. Participants who reported experiencing the visual illusion would show a disproportionately large number of random presses compared to those who did not.

Methods

Participants

We recruited 15 participants (all Leiden University students or recently graduated), out of a total of 24 that we aimed to test. There were two reasons for this reduced sample size. Firstly, due to COVID-19 related university closures and the need to finalize the project as soon as possible, we decided to test fewer participants to make this thesis possible. Secondly, it became apparent while running the study that the stimuli we used induced a visual illusion in some of the participants and as such it was decided to use this initial sample as a pilot. Out of the 15 people, 3 were excluded (pp 9, 12 and 15), due to the eye-tracking data exhibiting patterns indicative of an erroneous recording (large differences in pupil size between eyes, abnormal sudden dilation patterns). This resulted in a final N of 12 individuals, with a mean age M = 23.71 (SD = 3.33, min/max = 18/28, 4 male). Study exclusion criteria included any psychiatric disorders or physical disorders that could bias results, wearing contact lenses, and being younger than 18 years old or older than 30. Participants who wore glasses could participate after testing the eye-tracker calibration and confirming that proper recordings were being obtained. Approval for conducting the study was obtained from the Leiden University Ethics Committee. Participants signed a consent form prior to participating and received compensation at the end of the study (7.5 Euro/hour or course credits).

Task

We designed a custom dual modality (visual and auditory) oddball task, based partly on that used by Lakatos et al. (2016). Participants were required to respond to a rare stimulus in a cued modality by pressing the key “B” on a keyboard and to otherwise withhold any response. The task contained 10 blocks (with breaks in between) of 605 trials each (visual and auditory combined), in which participants were presented with simultaneous visual and auditory streams of cues but only had to attend to one of the modalities (they were prompted to ATTEND VISUAL or ATTEND AUDITORY at the beginning of each block) and to respond to the oddball in that modality. Half of the blocks required the participant to attend to the auditory stimuli and half to the visual. Only 15% of the total number of stimuli were targets. The auditory stimuli consisted of short beeps with a 460 Hz pitch for the standard cue and 476 Hz pitch for the target and the visual stimuli consisted of colored isoluminant RGB Gabor patches with a 10 degree orientation and a frequency of 0.1 cycles/pixel, created using the Gabor generator provided by Mathôt, (n.d).. The visual cues were presented at a frequency of 1.1 Hz and the auditory at 1.4 Hz. We chose standard and target stimuli that were very similar to each other to ensure that the stimulus presentation was near the detection threshold, thus reducing the chance of strong stimulus-evoked EEG and pupil responses. The average inter-stimulus interval (ISI) was 387 ms, with a range from 0 to 727 ms.

Procedure

Participants arrived during their allotted time and were presented with all information relevant for their participation. They were informed about potential risks and were asked again whether they were physically and psychologically fit to participate. If everything was in order, they were encouraged to ask questions at any point throughout the study and were prompted to sign a consent form. Afterwards they were asked to fill in a brief screening questionnaire, comprised of several common demographic questions as well as questions about their health and were then directed to the testing room where they were fitted with an EEG device, a head rest, and given a set of headphones to use throughout the study. I will not be discussing EEG results in this thesis, but it is important to be mindful of the potential effect that EEG preparation may have had on participants' mood or other aspects that could impact pupil recordings.

After the EEG setup was completed, participants filled in a mood and affect grid and were then instructed to stay as still as possible throughout the blocks and to always look at the middle of the screen, as to avoid artifacts in the recording. Participants were familiarized with the task by performing two training blocks (one for each attending condition), in which their performance was not recorded. Each block lasted 2 minutes and was played in real speed, containing 256 stimuli (visual and auditory combined). The instructions for the study were presented in writing on the task computer and the participant could read them before the experimenter repeated them out loud and asked if there were any questions. The instructions included general information about the study as well as side-by-side images of the visual stimuli, with explanations of what the participant had to do when viewing them. Because of the speed of the presentation, the responses could not be fast enough to fall right on the target and as such other stimuli could have overlapped with the responses. The participants were told not to be concerned with this and not to worry if they felt like they missed the target but rather to just do their best to respond as fast as possible. The experimenter also exemplified the audio stimuli using a phone (Samsung Galaxy S8), by playing the sounds in succession and asked the participant if they can differentiate them. After confirming that they understood the difference, participants' chins were positioned in the chin rest, the monitor was set up at 73 cm from their eyes and they were told to perform the practice trials. After the practice trials, the eye-tracker was calibrated and the participants performed the task, alone, in a dim lit room. Between blocks, participants could take breaks and decide when to start the following block. The task itself was rather simple and no participants appeared to struggle with the task, nor did they ask for clarifications after starting the task. Participants took between 60 and 70 minutes to complete the study (50 minutes task + breaks). The total length of the experiment was about 2 hours from beginning to end, depending on how fast preparations could be done.

Design

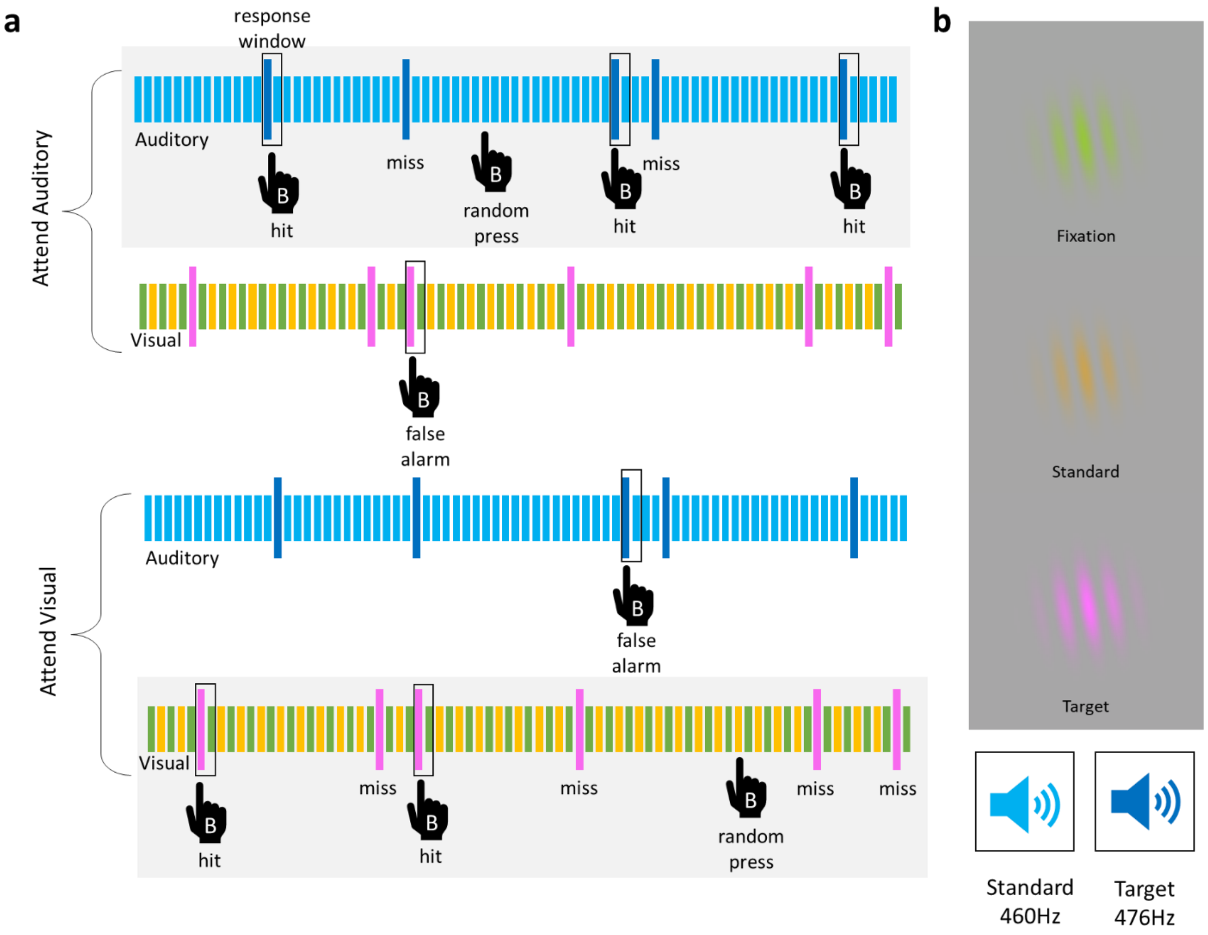
We performed our study within subjects, around our fixed factor which was the stimulus stream (combined auditory and visual). The number of hits, misses, false alarms, random presses and the pupil size constituted our dependent variables, and the order of the trials was pseudo-randomized with all subjects performing the task with the same pre-defined order. Figure 1 shows our study design and the colors of the visual stimuli used, as well as the pitch frequency of our auditory stimuli.

Pupillometry

The study took place in a dimly lit room while the participants kept their head in place using a chin rest. Participants were instructed to always look in the center of the screen, regardless of the attending condition they were in. The pupil diameter of both eyes was recorded using a sampling rate of 120 Hz and the calibration was performed once, at the beginning of the task, using a 5-point fixation procedure. The Tobii device automatically marked periods when no pupil signal was recorded (blinks) and custom code was compiled to save accurate timing data that could then be used to create eye tracking epochs and segmentations for the behavioral data. Additional custom MATLAB code was written to create PhysioData files for the pupil data and to derive the proportion of hits, misses and false alarms and the count of random presses for the behavioral data.

Figure 1

Multimodal oddball task design and stimuli



Note. (a) The study consisted of 10 blocks (5 auditory) in which participants were prompted to attend to one of the two modalities and respond to a rare stimulus. The average ISI was 387ms and the stimulus presentation frequency was 1.1 Hz for the visual stimulus and 1.4 Hz for the auditory. Key presses that fell on a target stimulus or on the stimulus immediately following it were marked as hits. Key presses falling on or immediately after a target in the non-attended modality were marked as false alarms. Any key presses that did not correspond with target stimuli were registered as random presses. (b) Visual stimuli used in the task and the pitch frequency of the auditory stimuli.

Behavioral data

Due to the rapid succession of stimuli which was faster than normal reaction times, the percentage of hits and false alarms was computed by including the key responses that fell within the same target trial or the one immediately following it. Hits represent the correct responses to the attended modality while false alarms represent responses to the targets of the non-attended modality divided by the total number of trials in that non attended modality. The percentage of misses was calculated as the total number of targets in the attended modality minus the hits. The number of random presses included all the key presses that were neither hits nor false alarms (so responses to non-targets). We also measured the participants' mood and arousal levels at the start and end of the task.

The Visual Illusion

An unforeseen factor that affected our study was the reporting of a visual illusion that was experienced by 5 participants (33%), 4 of which were included in the final analysis sample. It is unclear whether other participants also experienced the illusion but did not report it. Our visual stimuli were green (fixation), orange (frequent stimulus) and magenta (infrequent target stimulus). The illusion in question concerned complementary afterimages, a phenomenon by which a complementary colored image persists in the viewer's vision after the original image they viewed disappeared (Manzotti, 2017). These afterimages usually appear after prolonged exposure to the original images and are generally more salient than positive afterimages, which retain the original colors of the stimulus but may sometimes only last as little as 500 milliseconds (Cohen-Duwek & Spitzer, 2018). Examples of complementary afterimages can be found on the webpage of UNSW School of Physics,(n.d.). Although concern for potential conflicting visual effects of the stimuli existed during the design of the study, complementary afterimages were not considered, given that they typically require an exposure to the original image of at least 20-30 seconds (DeValois & Webster, 2011), whereas our stimuli appeared in very fast succession. This illusion may have complicated task performance for the participants because the green Gabor patch was presented as a fixation before every standard and target stimulus and thus, participants who experienced the illusion may have perceived the magenta afterimage of the green pattern very frequently and confuse it with the target. It is unclear whether participants could differentiate between the two magenta hues but some of them did mention that they could distinguish them (the target appeared slightly darker).

It must also be mentioned that 3 participants performed the task with different colors (blue for fixation, yellow for standard and orange for target) but it was decided to return to the original colors for consistency. The data of these three participants did not appear to differ from the other collected data in terms of the target measurements so it was also included in the analysis. Although it is unclear why the illusion occurred given our study design, we believe that it might have to do with how much the participants focused on the stimuli, as some reported to only see the illusion when focusing very hard on the visual stimuli. Moreover, despite the fact that traditionally it is believed that these negative afterimages are mostly generated in the primary visual pathway as a result of the overstimulation of retinal photoreceptors, some researchers have suggested that top-down processing (via priming) may also influence when the illusion manifests (Utz & Carbon, 2015). As such it is plausible that being instructed to attend to the stimuli and becoming primed to the patterns may have increased the likelihood that participants would experience the illusion. Unfortunately, the size of our sample as well as the disproportionately small number of participants who reported the illusion compared to those who did not, did not allow us to run statistical checks on this effect. Instead, we relied on comparing the average number of random presses of those who reported the illusion compared to the ones who did not and found both values to be very low. We expected that participants who reported the illusion would exhibit unusually large random press scores as they would basically confound the target with the standard cue. This however did not appear to be the case (see results section), so we decided to include their data in the main analysis.

Equipment and Programs

Participants performed the task using a research desktop computer with a refresh rate of 60.015 Hz and the sound volume set on 15. Pupil diameter measurements were acquired using the Tobii Pro X3-120 EPU (Tobii AB, 2020) eye tracker with a sampling frequency of 120Hz and a 5 point calibration protocol. The task was programmed using Eprime 3.0 with the appropriate Tobii extensions installed. For the auditory stimuli custom air headphones for in-ear use designed by The Support for Research, Laboratories and Education (SOLO) department at Leiden University were used due to their lack of wiring which prevented interference with the EEG setup. We acquired the EEG data using a BioSemi ActiveTwo system with 32+6 active AG-AgCl electrodes, digitized at 512 Hz. The electrodes were arranged according to the international 10-20 system. In addition to the 32 main electrodes, two reference electrodes were positioned on the mastoids, two facial electrodes above and below the left eye and two electrodes on the temples. The pupil data was analyzed using the PhysioData Toolbox v0.5.0 (Sjak-Shie, E. E., 2019). The code needed for converting the gaze data into PhysioData compatible files and for extracting the behavioral data along with the corresponding markers was compiled using MATLAB R2018b. Data analysis was performed using R and SPSS.

Results

Data pre-processing

The eye-tracking data was pre-processed and divided into epochs using the pupil diameter module of the PhysioDataToolbox (v.0.5.0) program (Sjak-Shie, E. E., 2019) based on the guidelines provided by Kret and Sjak-Shie (2019), to ensure the appropriate rejection of poor signal instances. Each of the 10 blocks for each participant was divided into static time windows of 30 seconds. This resulted in 10 time-widows per block. We removed the first time window of each block to allow for the stabilization of the pupil signal after the break and also removed the last window of each block as it was much shorter than the others (between 5 and 15 seconds depending on the exact length of each block which varied based on the ISI). After removal we were left with 8 windows of 30 seconds per block, for a total of 80 windows per participant.

The behavioral data (hits, misses, false alarms and random presses) was pre-processed using a custom MATLAB script and compiled into an output consisting of proportions of hits, misses and false alarms and counts of random presses, for each time-window which was then merged with the pupil data consisting of the mean pupil diameter (as calculated per time-window from the valid sample signal of both pupils after interpolation and smoothing; see appendix D). Given that the proportion of hits and that of misses are perfectly negatively correlated, we decided that misses would not add any explanatory variance to our study and decided not to use them in any of our analyses.

Across all participants, 1.75% of the time windows contained missing values. In the hits, misses and false alarm columns, missing values were imputed with the mean of the block in which they occurred. Due to the skewed distribution of random press counts (80% of the windows had 0 random presses), the missing values in this variable were set to zero.

Normal distribution testing

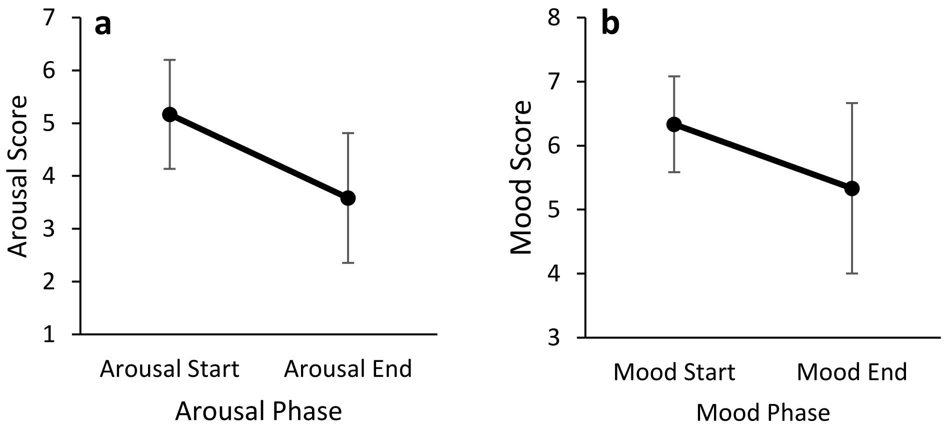
We conducted two Shaphiro-Wilk tests to explore whether our data was normally distributed. The first test was performed at the beginning of our analysis and included the data for each variable (pupil diameter, behavioral variables, mood and affect scores) across all 12 participants. We found that all our variables, except for mood and affect violated the assumption of normality. The second test was performed on the dataset containing only the pupil diameter, the hits, and the false alarms after it was processed using a median split based on pupil diameter and was separated according to the two attended modalities (visual and auditory) and revealed that all variables except for the proportion of false alarms were normally distributed. Detailed tests result for these checks can be found in Appendix A.

Mood and affect

We used two paired t-tests to examine whether the participants' level of arousal and their mood changed significantly throughout the study. As shown in Figure 2, both measures were lower at the end of the study compared to the beginning but only arousal ($t(11) = 2.77$, p = 0.018) appeared to be significantly lower whereas mood was not ($t(11) = 1.773$, p = 0.104).

Figure 2

Average difference in arousal and mood between the start and end of the study



Note. (a) Difference in self-reported arousal between the start and the end of the study, with error bars representing 95% confidence intervals (b) Difference in self-reported mood between the start and the end of the study, with error bars representing 95% confidence intervals

Random presses

We examined the number of random presses to investigate whether the visual illusion influenced participants' responses. Out of 12 included participants, 4 reported seeing the illusion. On average participants who reported the illusion made 0.26 random presses per minute while those who did not made 0.42 random presses per minute. Based on the distribution of random presses between the two groups (as shown in Figure 3), it appears that those who saw the illusion gave fewer random responses that those who did not. However, this difference cannot be reliably tested due to the small sample size and the scarcity of random presses (80% of time windows across all participants had 0 random presses). We decided not to exclude participants based on or to correct for this issue in our further analysis.

Figure 3

Average number of false alarms/minute for individuals who reported seeing the illusion and those who did not.

Note. Proportion of false alarms/minute made by participants who either did or did not report the illusion. The central line is the median, the x is the average, the boxes indicate the first and third quartile and the whiskers show the maxima and minima of the data.

Checking for time-on-task effects

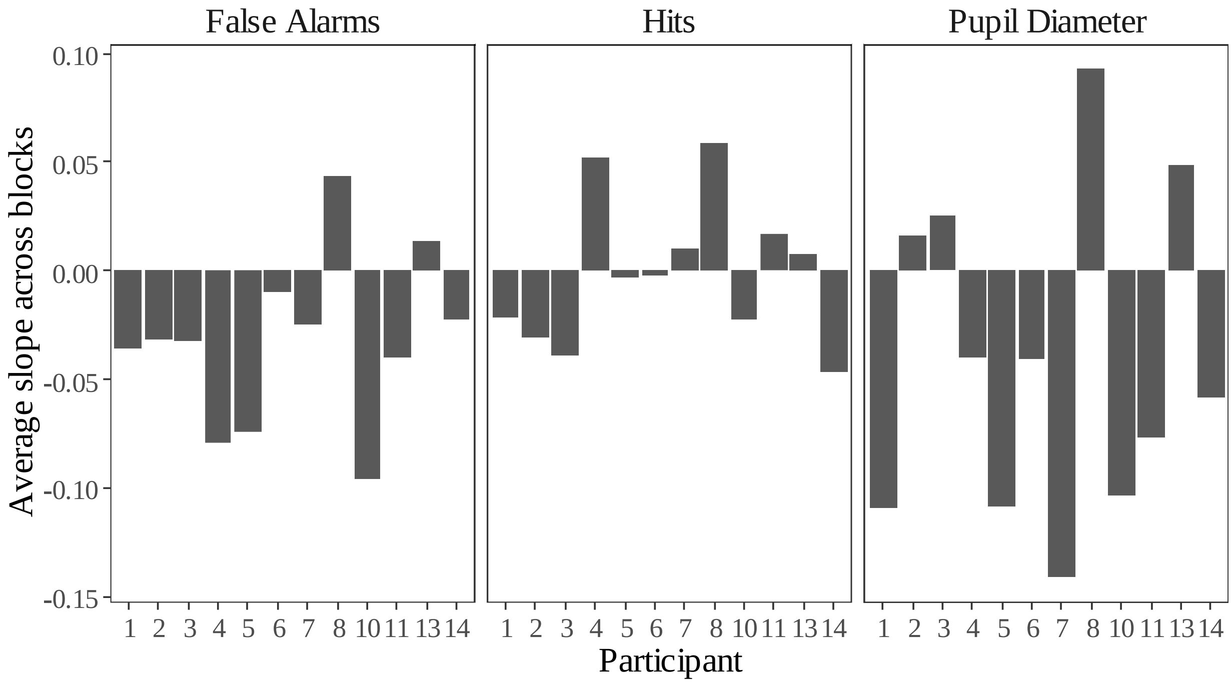
We continued our analysis by investigating whether our results were affected by time-on-task effects, that is to see whether performance degraded, and pupil diameter changed over the course of a block. For this part and all subsequent parts of the analysis we used the pupil size measurements, the proportion of hits and the proportion of false alarms.

Following Van Den Brink et al.'s (2016) model of testing for time-on-task effects, we transformed all our variables of interest into Z scores. While Z-scoring of proportions and variables with a high occurrence of zero values is problematic, it nevertheless was necessary to allow calculating the regression slopes needed for our analysis. After performing the Z-scoring, we fitted a straight line to each variable, for each block per participant and extracted the slopes of these lines using the *lm* (linear model) function in R. We then averaged these slopes across blocks for each participant and performed a one tailed t-test to check if the distribution of slopes was significantly different from 0, thus checking if they were on average decreasing, increasing, or staying the same. Figure 4 shows the slope trends for each of our three variables of interest.

Our results showed that the proportion of false alarms appeared to significantly decrease over the course of the blocks *t*(11) = -2.889, p = .015 while neither the pupil diameter (*t*(11) = -1.968, p = .075) nor the proportion of hits (*t*(11) = -.183, p = .858) showed any significant change over time. No false alarms were recorded in 94% of all time windows.

Figure 4

Slope trends for each variable of interest



Note. Average slope trend per behavioral measure, for each participant. The direction of most of the bars indicates whether, on average, that timeseries was increasing or decreasing.

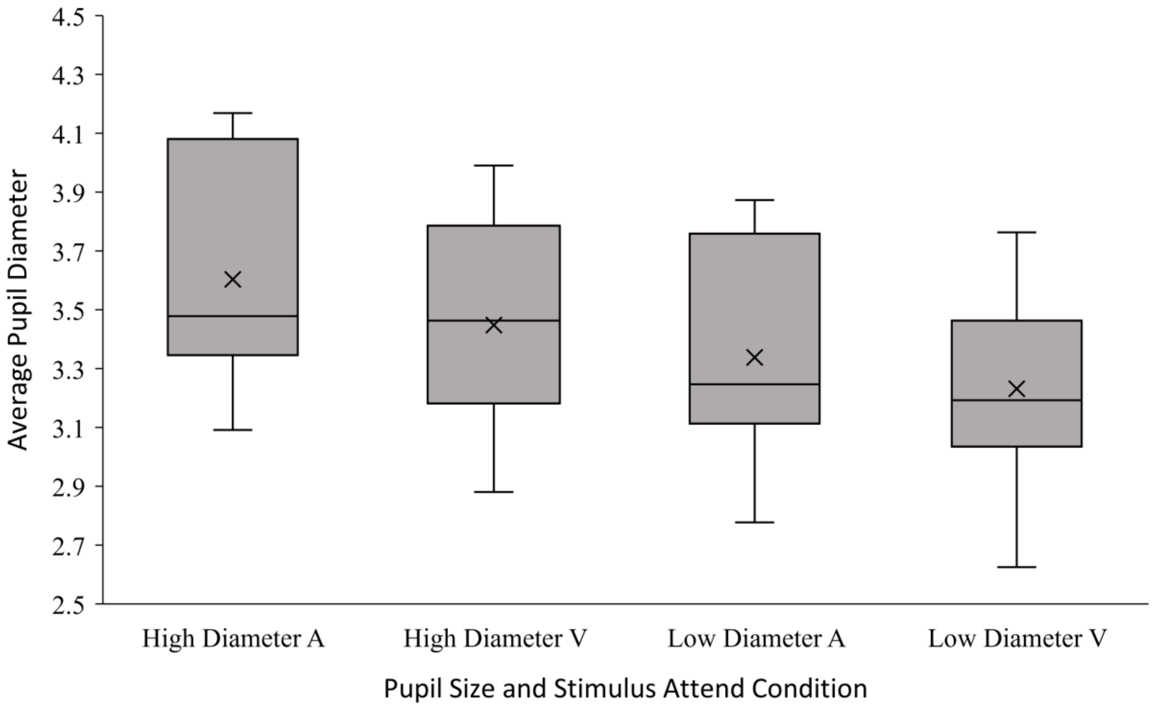
Performance differences based on pupil size

After exploring the effect of time-on-task we performed a median split of hits and false alarms based on pupil size within each block for every participant, to investigate whether performance was different when the pupil size was larger compared to when it was smaller. This resulted in four low-mean pupil and four high-mean pupil time windows in each block. We further divided our data based on the stimulus attend condition (auditory vs visual) to also examine whether performance differed based on this factor. Because our data was found to be normally distributed, we conducted several repeated measures ANOVAs to test our predictions.

We found a significant main effect of stimulus type on pupil diameter *F*(1,11) = 11.8, p = .006, with larger pupil sizes during the attend auditory blocks than during the attend visual blocks in both the low (*t*(11) = 2.55, p = 0.27; M = 3.34; SD = .36) and high (*t*(11) = 4.32, p < .001; M = 3.6; SD = .39) mean pupil diameter condition (low M = 3.2, SD = .34; high M = 3.4, SD = .36). These differences between conditions are exemplified in Figure 5. For the proportion of hits, we found no effect of stimulus type, *F*(1,11) = .009, p = .926, nor of pupil size condition, *F*(1,11) = .971, p = .346, suggesting that at different pupil sizes, participants did not perform the task with markedly different accuracy. Similarly, we found no effect of attend modality, *F*(1,11) = .402, p = .539, nor of pupil size, *F*(1,11) = .1.674, p = .222, on the proportion of false alarms.

Figure 5

Average pupil diameter in each of the attending conditions



Note. Average pupil diameter for each of the stimulus types. The central line is the median, the x is the average, the boxes indicate the first and third quartile and the whiskers show the maxima and minima of the data.

Discussion

We used a high-speed multimodal oddball task to investigate whether lapses in attention correspond to differences in pupil diameter. Based on previous research indicating vastly different pupil outcomes for attentional lapses (small, large, both small and large) we decided to approach the question in an exploratory manner and thus did not formulate a direction for the effect of pupil size. It must be noted that due to the small size of our sample and to the fact that the data could not be analyzed in an ideal manner, I will discuss the results both based on the significance testing as well as general trends observed.

Arousal, more so than mood, has been shown to play an important role in attention by enhancing the processing of task-relevant salient stimuli (Sutherland & Mather, 2018), with increased (but not extreme) levels of arousal improving attentional performance. Because of this we designed the study in such a way that it would not be long enough to completely exhaust the participants, but still long enough to ensure a certain degree of fatigue and reduction in arousal, thus increasing the likelihood of attentional lapses. We found that throughout the study both the participants' mood and arousal levels decreased (albeit only the decrease in arousal was significant), which only partially confirms our hypothesis that the two measures would significantly reduce over time. The reduction in reported arousal between the start and the end of the study indicates that our task length manipulation did work as expected. Nevertheless, as we will be discussing further, this does not appear to have translated into attentional lapses.

We further examined the number of random presses and hypothesized that if the visual illusion affected participants' performance, they would show abnormally large scores of random presses because standard stimuli in the visual attend condition would also be signaled as targets due to them appearing magenta. We used this measure instead of the number of hits because the visual illusion presented itself in the same color as our targets and as such, we had no reason to believe that participants' performance would be affected on the target trials. While our dataset was too small for running a formal statistical test, we found that, on average, participants made fewer than one random press per minute, regardless of whether they reported the illusion. In contrast to our expectations, our results suggest that individuals who saw the illusion in fact performed fewer random presses than those who did not, but it is unclear whether this was due to an actual difference in performance or because of the small number of people who reported it. A possible interpretation of this finding is that experiencing the illusion caused the participants to focus more on the task due to the increased difficulty of differentiating target from standard stimuli. If increased focus were the reason for this effect, then an increase in performance in the visual condition would likely also be observed. While our sample did not allow for meaningful analysis in this respect, this effect may be investigated in the future. Another explanation for the low number of random presses in affected participants could be that they could nonetheless differentiate between the target color and that of the visual illusion. In fact, two individuals did report at the end of the study that the target appeared darker than the illusion. Altogether, while we decided based on the small number of random presses not to exclude participants who reported the illusion, we will nevertheless remain mindful of their potential confounding effect on our dataset when drawing further conclusions from our study.

Similarly to Van Den Brink et al. (2016), we investigated whether our variables of interest were affected by time-on-task effects. Our findings contrast those of the previously mentioned researchers in that we did not find any such effects for our variables, except for the false alarms which appeared to decrease over time. It must be noted that although false alarms are generally considered to be indicative of lapses of attention in continuous performance tasks

(Esterman et al., 2013), previous studies examining attention and entrainment in a dual modality oddball task (Lakatos et al., 2016) suggest that false alarms are merely indicators of a switch of attention toward the irrelevant perceptual modality, whereas overall decreases in both hits and false alarms would suggest genuine attentional lapses. Given that in our study a reduction in false alarms did not correspond with a reduction in hits, we cannot conclude that this is indicative of lapses in attention. Instead, the decrease in false alarms, could be explained by the fact that, participants' task performance may have needed some time to stabilize after each break between blocks and that they required some time to re-engage with the task and focus. Nevertheless, our initial hypothesis was that the proportion of false alarms would increase over time in each block, as participants became fatigued, bored, and prone to disengage from the task. We can only assume that while participants did indeed show decreased levels of arousal, the task might have still been easy enough that they learned how to perform it efficiently.

A temporal trend of decrease in pupil diameter was also observed. While not significant at an alpha = .05 level, the pupil diameter scores showed significant decreases over time at an alpha =.1 level. Based on previous pupil studies that employed larger samples we have reasons to believe that this relationship between time elapsed and pupil diameter would likely also become significant, indicating that the pupil too may be affected by time-on-task effects. Interestingly, the proportion of hits did not appear to show a decrease over time even though performance degradation is common in tasks that are highly demanding and lengthy. Much like the low proportion of false alarms this might indicate that the task may have been relatively easy to perform.

Previous studies found that pupil diameter is closely related to task performance in attention tasks, although a general lack of consensus exist as to whether smaller or larger pupil diameter (or indeed both) indicate lapses in attention. We found no effect of pupil size on our behavioral measures of interest (hits and false alarms) which at first glance appears to contradict previous research. It could, however, be the case that our analysis was not suitable for picking up on subtle relationships that could easily be obscured by latent time-on-task effects or suboptimal pupil measurements. Van Den Brink et al. (2016) suggested that the pupil derivative, consisting of the average difference between consecutive samples, may be a better measure of pupil change than diameter per se. These results do not confirm our hypothesis that performance would be different depending on pupil diameter, but further investigation is necessary.

Although task performance did not appear to be related to pupil size, we did find a highly significant effect of the attended modality on pupil diameter (despite identical visual input, assuming that participants fixated the center of the screen), with larger average diameters in the auditory attend condition than in the visual attend blocks. We did not find an effect of attended modality on our behavioral measures, however. Previous studies have found that differences do exist between modalities in terms of entrainment to task stimuli (VanRullen et al., 2014) and it is plausible that these neurally driven differences could also manifest as changes in pupil diameter. The fact that we only found differences between stimulus types in the pupil diameter and not in the behavioral data may suggest that different task modalities may require different amounts of attentional focus to ensure optimal performance. The two modalities may not have been precisely matched in difficulty, leading to the visual task being easier to perform than the auditory, thus incurring less mental load and smaller pupil sizes.

To summarize, the current pilot study partially confirms previous findings regarding pupil diameter and attentional lapses, but due to the reduced sample size these results cannot be generalized. Nevertheless, some of the observed trends align well with existing literature and should be followed up on with a more powerful sample and an improved methodology within the same paradigm. We found that participants' self-reported levels of arousal decreased within a block over the course of the task. This fits well with previous research showing that lengthy and attention-intense studies cause fatigue and reduced arousal levels. We also found a decrease in false alarms which may indicate that participants' performance was poorer at the beginning of the block as they re-engaged with the task after the break and resumed attentional focus. Given that this decrease in false alarms was not accompanied by a decrease in hits, we cannot say that participants paid less attention to the task as the study progressed. Despite our expectation that different pupil sizes would correspond to markedly different behavioral measures, we did not find this to be the case, which appears to contradict previous literature. The reason for this could be our sample: Drawing conclusions about attentional lapses based on our very limited number of participants is difficult (if not impossible), but this study offers important lessons for future similar pieces of research. Although the visual illusion did not appear to have affected the number of false alarms participants made, a significant difference in pupil size between the visual and the auditory conditions was found, suggesting that either our two stimulus types were not well matched in terms of difficulty or that the illusion influenced pupil diameter without influencing behavior. To our knowledge, this kind of interference from a complementary after-image illusion has not been reported in any previous attention studies using a similar design paradigm as ours. This makes it an interesting effect to explore in more detail and an important methodological consideration for future research.

References

Aston-Jones, G., & Cohen, J. D. (2005). Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *Journal of Comparative Neurology*, *493*(1), 99–110. https://doi.org/10.1002/cne.20723

Cohen-Duwek, H., & Spitzer, H. (2018). A Model for a Filling-in Process Triggered by Edges Predicts "Conflicting" Afterimage Effects. *Frontiers in Neuroscience*, *12*(AUG), 559. https://doi.org/10.3389/fnins.2018.00559

Coull, J. T., Jones, M. E. P., Egan, T. D., Frith, C. D., & Maze, M. (2004). Attentional effects of noradrenaline vary with arousal level: Selective activation of thalamic pulvinar in humans. *NeuroImage*, *22*(1), 315–322. https://doi.org/10.1016/j.neuroimage.2003.12.022

DeValois, K., & Webster, M. (2011). Color vision. *Scholarpedia*, *6*(4), 3073. https://doi.org/10.4249/scholarpedia.3073

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. In *Neuroscience and Biobehavioral Reviews* (Vol. 81, pp. 181–187). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2017.02.011

Dugué, L., Marque, P., & VanRullen, R. (2015). Theta oscillations modulate attentional search performance periodically. *Journal of Cognitive Neuroscience*, *27*(5), 945–958. https://doi.org/10.1162/jocn_a_00755

Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the Zone or Zoning Out? Tracking Behavioral and Neural Fluctuations During Sustained Attention. *Cerebral Cortex*, *23*(11), 2712–2723. https://doi.org/10.1093/cercor/bhs261

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*(JUL), 154. https://doi.org/10.3389/fpsyg.2011.00154

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

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Publishing Group, 15*(4). https://doi.org/10.1038/nn.3063

Hafed, Z. M., Chen, C. Y., & Tian, X. (2015). Vision, perception, and attention through the lens of microsaccades: Mechanisms and implications. In *Frontiers in Systems Neuroscience* (Vol. 9, Issue DEC, p. 167). Frontiers Research Foundation. https://doi.org/10.3389/fnsys.2015.00167

Hanslmayr, S., Axmacher, N., & Inman, C. S. (2019). Modulating Human Memory via Entrainment of Brain Oscillations. In *Trends in Neurosciences* (Vol. 42, Issue 7, pp. 485–499). Elsevier Ltd. https://doi.org/10.1016/j.tins.2019.04.004

Helfrich, R. F., Breska, A., & Knight, R. T. (2019). Neural entrainment and network resonance in support of top-down guided attention. In *Current Opinion in Psychology* (Vol. 29, pp. 82–89). Elsevier B.V. https://doi.org/10.1016/j.copsyc.2018.12.016

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–1596. https://doi.org/10.1162/jocn.2010.21548

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

Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior Research Methods*, *51*(3), 1336–1342. https://doi.org/10.3758/s13428-018-1075-y

Lakatos, P., Barczak, A., Neymotin, S. A., McGinnis, T., Ross, D., Javitt, D. C., & O'Connell, M. N. (2016). Global dynamics of selective attention and its lapses in primary auditory cortex. *Nature Neuroscience*, *19*(12), 1707–1717. https://doi.org/10.1038/nn.4386

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. *Current Biology*, *29*(18), R890–R905. https://doi.org/10.1016/j.cub.2019.07.075

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, *22*(11), 1000–1004. https://doi.org/10.1016/j.cub.2012.03.054

Manzotti, R. (2017). A Perception-Based Model of Complementary Afterimages. *SAGE Open*, *7*(1), 215824401668247. https://doi.org/10.1177/2158244016682478

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: Prestimulus α phase predicts visual awareness. *Journal of Neuroscience*, *29*(9), 2725–2732. https://doi.org/10.1523/JNEUROSCI.3963-08.2009

Mathôt, S. (n.d.). *CogSci Gabor Generator*.

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154. https://doi.org/10.1002/hbm.22466

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1532–1543. https://doi.org/10.1111/j.1469-8986.2011.01226.x

Novembre, G., Sammler, D., & Keller, P. E. (2016). Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. *Neuropsychologia*, *89*, 414–425. https://doi.org/10.1016/j.neuropsychologia.2016.07.027

Ohki, T., & Takei, Y. (2018). Neural mechanisms of mental schema: a triplet of delta, low beta/spindle and ripple oscillations. In *European Journal of Neuroscience* (Vol. 48, Issue 7, pp. 2416–2430). Blackwell Publishing Ltd. https://doi.org/10.1111/ejn.13844

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. *Eye Tracking Research and Applications Symposium (ETRA)*, 141–144. https://doi.org/10.1145/1743666.1743701

Raichle, M. E. (2010). Two views of brain function. In *Trends in Cognitive Sciences* (Vol. 14, Issue 4, pp. 180–190). Elsevier Current Trends. https://doi.org/10.1016/j.tics.2010.01.008

Riecke, L., Formisano, E., Sorger, B., Başkent, D., & Gaudrain, E. (2018). Neural Entrainment to Speech Modulates Speech Intelligibility. *Current Biology*, *28*(2), 161-169.e5. https://doi.org/10.1016/j.cub.2017.11.033

Sara, S. J., & Bouret, S. (2012). Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. In *Neuron* (Vol. 76, Issue 1, pp. 130–141). Cell Press. https://doi.org/10.1016/j.neuron.2012.09.011

Smith, A., & Nutt, D. (1996). Noradrenaline and attention lapses. *Nature*, *380*(6572), 291. https://doi.org/10.1038/380291a0

Smith, A. P., Wilson, S. J., Glue, P., & Nutt, D. J. (1992). The effects and after effects of the α2-adrenoceptor antagonist idazoxan on mood, memory and attention in normal volunteers. *Journal of Psychopharmacology (Oxford, England)*, *6*(3), 376–381. https://doi.org/10.1177/026988119200600306

Sutherland, M. R., & Mather, M. (2018). Arousal (but not valence) amplifies the impact of salience. *Cognition and Emotion*, *32*(3), 616–622. https://doi.org/10.1080/02699931.2017.1330189

Tobii AB. (2020). *Tobii Pro X3-120 screen-based eye tracker*. https://www.tobiipro.com/product-listing/tobii-pro-x3-120/

UNSW School of Physics. (n.d.). *Complementary colours, after-images, retinal fatigue, colour mixing and contrast sensitivity*. Retrieved December 31, 2020, from https://www.animations.physics.unsw.edu.au/jw/light/complementary-colours.htm

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective and Behavioral Neuroscience*, *16*(4), 601–615. https://doi.org/10.3758/s13415-016-0417-4

Utz, S., & Carbon, C. C. (2015). Afterimages are biased by top-down information. *Perception*, *44*(11), 1263–1274. https://doi.org/10.1177/0301006615596900

Van Den Brink, R. L., Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil diameter tracks lapses of attention. *PLoS ONE*, *11*(10), 1–16. https://doi.org/10.1371/journal.pone.0165274

Van Orden, K. F., Jung, T. P., & Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology*, *52*(3), 221–240. https://doi.org/10.1016/S0301-0511(99)00043-5

van Wassenhove, V. (2016). Temporal cognition and neural oscillations. In *Current Opinion in Behavioral Sciences* (Vol. 8, pp. 124–130). Elsevier Ltd. https://doi.org/10.1016/j.cobeha.2016.02.012

VanRullen, R., Reddy, L., & Koch, C. (2006). The continuous wagon wheel illusion is associated with changes in electroencephalogram power at ∼13 Hz. *Journal of*

Neuroscience, *26*(2), 502–507. https://doi.org/10.1523/JNEUROSCI.4654-05.2006

VanRullen, R., Zoefel, B., & Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1641). https://doi.org/10.1098/rstb.2013.0214

Voloh, B., & Womelsdorf, T. (2016). A role of phase-resetting in coordinating large scale neural networks during attention and goal-directed behavior. In *Frontiers in Systems Neuroscience* (Vol. 10, Issue MAR, p. 18). Frontiers Media S.A. https://doi.org/10.3389/fnsys.2016.00018

Wass, S. V., Whitehorn, M., Marriott Haresign, I., Phillips, E., & Leong, V. (2020). Interpersonal Neural Entrainment during Early Social Interaction. In *Trends in Cognitive Sciences* (Vol. 24, Issue 4, pp. 329–342). Elsevier Ltd. https://doi.org/10.1016/j.tics.2020.01.006

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Appendix A

Normality Test Tables

Table A1

Results of the first Shapiro-Wilk test for determining whether the variables are normally distributed.

Variable	Statistic	df	Sig.
Mean diameter both eyes (mm)	.983	960	< .001
hits (%)	.836	960	< .001
Misses (%)	.836	960	< .001
False alarms (%)	.836	960	< .001
Random presses (count)	.404	960	< .001
Mood (Likert Score)	.928	11	.364
Arousal (Likert Score)	.869	11	.064

Note: The information in brackets represents the type of score of each variable.

Table A2

Results of the Shapiro-Wilk test for determining whether the variables are normally distributed, after the median split.

Variable	Statistic	df	Sig.
Large Pupil Diameter (LPD)	.933	12	.413
Small Pupil Diameter (SPD)	.958	12	.756
LPD Hits	.913	12	.236
SPD Hits	.917	12	.262
LPD False Alarms	.546	12	.000
SPD False Alarms	.733	12	.002

Note: LPD = Large Pupil Diameter; SPD = Small Pupil Diameter

Appendix B

Table with the number of random presses/minute

Table B1

Average number of random presses per minute, per participant, according to whether they reported the visual illusion or not, together with the average number of random presses per minute for the two groups (reported/did not report the illusion) separately.

PP	random presses/min	reported illusion
1	0.52	yes
2	0.32	no
3	0.1	yes
4	1.2	no
5	0.72	no
6	0.54	no
7	0.04	yes
8	1.06	no
10	0.16	no
11	0.16	no
13	0.14	no
14	0.22	yes
Average RP, Illusion Yes: 0.26		
Average RP, Illusion No: 0.42		

Appendix C

SPSS Syntax

#Normality test general dataset

```
EXAMINE VARIABLES=meanDia_BothEyes hits misses falsealarm
  /PLOT BOXPLOT STEMLEAF NPPLOT
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

 # Split file by participant; investigated the descriptive statistics; saved Z-scores

```
SORT CASES  BY PP.
SPLIT FILE SEPARATE BY PP.
DESCRIPTIVES VARIABLES=meanDia_BothEyes hits misses falsealarm
  /SAVE
  /STATISTICS=MEAN STDDEV MIN MAX SKEWNESS.
```

 # Unsplit file; performed t-test on average slopes for the pupil diameter, hits, and false alarms

```
SPLIT FILE OFF.
GET
  FILE='C:\Users\alfea\Desktop\analyses\use this data\final
version\slopes\avg_slope_stats.sav'.
DATASET NAME DataSet2 WINDOW=FRONT.
DATASET ACTIVATE DataSet2.
```

```
SAVE OUTFILE='C:\Users\alfea\Desktop\analyses\use this
data\final '+
    'version\slopes\avg_slope_stats.sav'
  /COMPRESSED.
T-TEST
  /TESTVAL=0
  /MISSING=ANALYSIS
  /VARIABLES=avg_slope_dia avg_slope_hits avg_slopes_fa
  /ES DISPLAY(TRUE)
  /CRITERIA=CI(.95).
```

 # Sorted data in ascending order by participant, block and mean pupil diameter

```
GET
  FILE='C:\Users\alfea\Desktop\analyses\use this data\final
version\sort_ascending.sav'.
DATASET NAME DataSet3 WINDOW=FRONT.
SORT CASES BY PP(A) block(A).
SORT CASES BY PP(A) block(A) meanDia_BothEyes(A).
DATASET ACTIVATE DataSet3.

SAVE OUTFILE='C:\Users\alfea\Desktop\analyses\use this
data\final version\sort_ascending.sav'
  /COMPRESSED.
DATASET ACTIVATE DataSet1.
DATASET CLOSE DataSet3.
DATASET ACTIVATE DataSet1.

SAVE OUTFILE='C:\Users\alfea\Desktop\analyses\use this
data\final version\beh_pupil_merged.sav'
```

```
  /COMPRESSED.
DATASET ACTIVATE DataSet1.
DATASET CLOSE DataSet2.
GET
  FILE='C:\Users\alfea\Desktop\analyses\use this
data\median_split_avg.sav'.
DATASET NAME DataSet4 WINDOW=FRONT.
DATASET ACTIVATE DataSet1.
NEW FILE.
DATASET NAME DataSet5 WINDOW=FRONT.

SAVE OUTFILE='C:\Users\alfea\Desktop\analyses\use this
data\final version\avg_high_vs_low.sav'
  /COMPRESSED.
```

 # Normality test for data split based on pupil size

```
EXAMINE VARIABLES=high_dia low_dia high_hit low_hit high_fa
low_fa
  /PLOT BOXPLOT STEMLEAF NPPLOT
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

 # Repeated measures ANOVAs to examine the relationship between variables based on pupil size and stimulus attend condition

```
DATASET ACTIVATE DataSet5.

SAVE OUTFILE='C:\Users\alfea\Desktop\analyses\use this
data\final version\avg_high_vs_low.sav'
```

```
  /COMPRESSED.
GLM low_hits_A low_hits_V high_hits_A high_hits_V
  /WSFACTOR=low_high_comp 2 Polynomial a_v_comp 2 Polynomial
  /METHOD=SSTYPE(3)
  /PRINT=DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /WSDESIGN=low_high_comp a_v_comp low_high_comp*a_v_comp.
GLM low_fa_A low_fa_V high_fa_A high_fa_V
  /WSFACTOR=low_high_comp 2 Polynomial a_v_comp 2 Polynomial
  /METHOD=SSTYPE(3)
  /PRINT=DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /WSDESIGN=low_high_comp a_v_comp low_high_comp*a_v_comp.
GLM low_dia_A low_dia_V high_dia_A high_dia_V
  /WSFACTOR=low_high_comp 2 Polynomial a_v_comp 2 Polynomial
  /METHOD=SSTYPE(3)
  /PRINT=DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /WSDESIGN=low_high_comp a_v_comp low_high_comp*a_v_comp.
```

 # Descriptive statistics for pupil diameter in the two stimulus attend conditions and pupil sizes

```
DESCRIPTIVES VARIABLES=low_dia_A high_dia_A low_dia_V high_dia_V
  /STATISTICS=MEAN STDDEV MIN MAX.



DATASET ACTIVATE DataSet5.

SAVE OUTFILE='C:\Users\alfea\Desktop\analyses\use this
data\final version\avg_high_vs_low.sav'
```

```
  /COMPRESSED.
DATASET ACTIVATE DataSet1.
GET
  FILE='C:\Users\alfea\Desktop\analyses\use this data\final
version\slopes\avg_slope_stats.sav'.
DATASET NAME DataSet6 WINDOW=FRONT.
DESCRIPTIVES VARIABLES=avg_slope_dia avg_slope_hits
avg_slope_misses avg_slopes_fa
  /STATISTICS=MEAN STDDEV MIN MAX.
```

#T-test to check whether the pupil diameter (large and small) in the auditory condition differed from that in the visual condition

```
T-TEST PAIRS=low_dia_A high_dia_A WITH low_dia_V high_dia_V
(PAIRED)
  /ES DISPLAY(TRUE) STANDARDIZER(SD)
  /CRITERIA=CI(.9500)
  /MISSING=ANALYSIS.
```

Appendix D

R Syntax

Part 1: Data Pre-processing

 # Behavioral data was arranged in the order it had to be in 'A1','V1','A2','A3','V2','A4','V3','V4','A5','V5'

```
library(dplyr)
library(tidyr)

Thesis_Analysis <-
read.csv("C:/Users/alfea/Desktop/analyses/Proto/Thesis
Analysis.csv")

names(Thesis_Analysis)[1] <- 'PP'

block_order_vector <-
c('A1','V1','A2','A3','V2','A4','V3','V4','A5','V5')

Thesis_Analysis <- Thesis_Analysis %>%
  mutate(cond_block = paste(Condition, Block, sep = '')) %>%
  mutate(block_order = 0)

border <- 1
for (current_condblock in block_order_vector) {
  Thesis_Analysis$block_order[which(Thesis_Analysis$cond_block
== current_condblock)] <- border
  border <- border + 1
}

thesis_analysis_sorted <- Thesis_Analysis %>%
```

```
  arrange(PP, block_order, Time_Window)

write.csv(thesis_analysis_sorted,
'thesis_timed_analysis_sorted_by_blockorder.csv')
```

Part 2: Extract pupil data

```
library(dplyr)
library(stringr) # Dealing with strings (regex)



physio_data_raw <- readxl::read_xlsx('D:/Google
Drive/Thesis/Gaze Data
Analysis/convertGazedataWithLabelsFromText/data/PhysioEpochResul
ts_29.12.2020.xlsx',
                                 skip = 4)

block_regex_pattern <- '(?<=(PupilAnalyzer_))[a-zA-Z0-9]+(?=[_
])'
time_window_regex_pattern <- '(?<=_)[0-9]+(?=( \\(Entrain))'
PP_regex_pattern <- '(?<=2020-)[\\d]{1,2}'



physio_data <- physio_data_raw %>%
  select(Epoch_ID, meanDia_BothEyes) %>%
  slice(1:1650) %>%
  mutate(block = str_extract(Epoch_ID, block_regex_pattern) %>%
           factor()) %>%
  mutate(time_window = str_extract(Epoch_ID,
time_window_regex_pattern)) %>%
  mutate(PP = str_extract(Epoch_ID, PP_regex_pattern) %>%
           factor()) %>%
```

```
  mutate(is_avg_row = is.na(time_window))

# Remove rows that only contain average data
physio_data_timed <- physio_data %>%
  filter(!is_avg_row) %>%
  mutate(block_order = 0, time_window = as.integer(time_window))

block_order_vector <-
c('Aud1','Vis1','Aud2','Aud3','Vis2','Aud4','Vis3','Vis4','Aud5'
,'Vis5')

border <- 1
for (current_condblock in block_order_vector) {
  physio_data_timed$block_order[which(physio_data_timed$block ==
current_condblock)] <- border
  border <- border + 1
}

excluded_PPs <- c(9, 12, 15)

physio_data_timed_ordered <- physio_data_timed %>%
  arrange(as.integer(as.character(PP)), block_order,
time_window) %>%
  filter(time_window != 1 & time_window != 10, !(PP %in%
excluded_PPs))

write.csv(physio_data_timed_ordered, 'physio_data.csv')
```

Part 3. Merge Behavioral and Eye-Tracking Data

```
pupil_data <- read.csv('physio_data.csv') # This is the final
data frame that is generated by extract_pupil_data.R
```

```
# This file is generated by matlab script
call_eprime_data_full_modified.m
behavioral_data_raw <- readxl::read_xlsx(r"(D:\Google
Drive\Thesis\Behavioral Data Analysis\Participants\Matlab
Behavior Analysis\behavior_results\timed_all_vars_1-15.xlsx)")



excluded_PPs <- c(9, 12, 15)

# Dplyr filter function returns what is LEFT after filtering

behavioral_data <- behavioral_data_raw %>%
  filter(time_window != 1) %>% # Filter out first time window
  filter(!(PP %in% excluded_PPs)) # Exclude participants

# ==== Prepare pupil data and behavioral data for merge ====
# Generate unique row IDs. row IDs consist of participant ID,
block, and time window, separated by underscore _
pupil_data <- pupil_data %>%
  mutate(row_id = paste(PP, block, time_window, sep = '_')) %>%
  select(row_id, meanDia_BothEyes)

behavioral_data <- behavioral_data %>%
  mutate(row_id = paste(PP, block, time_window, sep = '_'))

# Test whether row IDs are unique
testthat::expect_length(unique(behavioral_data$row_id),
nrow(behavioral_data))
testthat::expect_length(unique(pupil_data$row_id),
nrow(pupil_data))
```

```
# Merge
data_merged <- full_join(behavioral_data, pupil_data, by = 'row_id')

write.csv(data_merged, 'behavioral_pupil_data_merged.csv')
```

Part 4: Calculated the average slopes

```
setwd("C:/Users/alfea/Desktop/analyses/use this data/final version")



# Read in data from SORTED table
data_sorted <- foreign::read.spss("beh_pupil_merged.sav") %>% as.data.frame()

# Dplyr used for data wrangling
library(dplyr)

# for (current_PP in unique(data_sorted$PP)) {
#   regressions <- data_sorted %>%
#     select(PP == current_PP)
# }

# Calculate regression across all time windows per block per participant
regressions <- data_sorted %>%
  mutate(time_window_order = rep(rep(1:8, 10), 12)) %>% #creates the variable time_window_sorted, does not save it anywhere just creates it for this part
  group_by(PP, block) %>%
  group_map(~ broom::tidy(lm(Zfalsealarm ~ time_window_order, data = .x))) #group_map takes a function and applies it to each
```

```
group defined above .x means give it the whole dataset
data_sorted, .x is a parameter from group_map

# Extract slopes of all regressions from regression object list
slopes = c()
for (block in regressions) {
  slopes <- c(slopes, block$estimate[2])
}

# Generate vector containing participant numbers. Vector matches
length of number of regressions
# Does this by selecting all time windows 'S2', as they only
occur once per participant per block
PPs <- data_sorted %>%
  filter(time_window == '2') %>%
  select(PP)

# Combine participant IDs and slopes into data.frame to use for
per-ppt-averaging
PPs_slopes <- data.frame(PPs, slopes) %>%
  janitor::clean_names()

# Calculate average slopes across all blocks per participant
avg_slopes <- PPs_slopes %>%
  group_by(pp) %>%
  group_map(~ mean(.x$slopes)) %>%
  unlist()

# Save data to csv
write.csv(slopes, 'slopes_fa.csv')
write.csv(avg_slopes, 'avg_slopes_fa.csv')
print(paste('Files saved to', getwd()))
```