Pupillometry investigation in visual imagery extremes

Léo Pasturel[1,2,3*], **Nathalie Guyader**[1,4], **Alan Chauvin**[2,3]

***For correspondence:**
pasturelleo@gmail.com (LP)

[1]Grenoble Institute of Technology (G-INP); [2]University Grenoble Alpes (UGA); [3]Laboratory of Psychology and Neuro-Cognition (LPNC); [4]Grenoble Images Speech Signal and Control (GIPSA)

Abstract Visual mental imagery has been a central feature of cognitive processes for centuries. It has been interesting more and more researchers in order to understand its functions. Subjective measures were mostly used until recently when two researches found that pupil could be a good objective indicator of visual mental representations. We expected to replicate previous results which found that the pupil constrict while imaging bright stimuli and dilate while imaging dark stimuli. To do so, pupil size of participants from general population was recorded in two experiments. One that implicates visual imagery in short-term memory and one in long-term memory. Results were compared to a subjective scale. We did not find any link between pupil size during imagery and self reported visual imagery. Serendipitously, we found that the addition of dynamic variability to a given brightness level induces pupil constriction.

Word count: 7845

Introduction

A matter of imagery

Imagine sitting at your breakfast table. Here is your cup of coffee. On one side are some of your favourite fruits. On the other side, some crunchy toast, straight out of the toaster, with butter and marmalade. Your mouth starts to water at this scene as you take a bite, tasting the sweetness and richness while chewing. You can hear the sounds of eating. You put down your toast and grab your coffee, take the time to appreciate its aroma, and take a sip. Just as expected, its bitterness balanced perfectly with the sweetness of your toast.

What many of you just experienced is mental imagery. While reading the text above, most of your senses may have been stimulated: your proprioception, sight, touch, taste, smell and hearing. Most people will have felt the experience of sitting at their table, living this moment. However, do not assume that everyone felt the same as you did. In fact, some people would have experienced this vignette as if it were really before them, other would have read through the scene without any sensory experience and most people would sit on a continuum in between those extremes.

Historical context

According to ***Pearson et al.*** (***2015***), mental imagery refers to "representations and the accompanying experience of sensory information without a direct external stimulus". In other words, what we call mental imagery is a mental representation that does not require any sensory input to be triggered. In this study, we focus on *visual mental imagery* that is images that form in the mind's eye and do not require any visual stimulus. This mechanism has been intriguing researchers since the

ancient Greek philosophers Plato and Aristotle who coined the term "*phantasma*" (plural: "*phantasmata*") to refer to mental imagery. At this time, Aristotle wrote that it was "impossible to think without an image" ***Thomas*** (***2021***). For Aristotle, even if we think of something that has no quantity, or "spatial extension" (as extrapolated by Gregoric) there would always be something spatially determined before our mind's eye (***Gregoric, 2007***). This definition definitely places imagery at the center of any cognitive process. Afterwards, the topic was set aside for many centuries until the 19th century. The British statistician and psychologist Sir Francis Galton was one of the first to report research on the differences between people in their experience of visual imagery (***Galton, 1880***). To do so, he created the "breakfast-table" questionnaire. There, he asked his participants to recall their breakfast table from the morning and to consider the picture that comes to their mind's eye. This questionnaire focuses on several points. The illumination of visual imagery, to determine if it is dim or clear, if the brightness is comparable to the one of the actual scene. The definition of the image, to determine if the whole scene is perceived clearly at the same time or if only one point of interest is sharp at a time. The colours of the image, to determine whether or not colours are distinguishable. The questionnaire was given to 272 participants, all men. There, Galton reported large differences between people going on a continuum from a complete lack of visual imagery to a life-like visual imagery. He even found that scientific men, as a class, have feeble power of visual representation. His conclusion is that thinking hard on generalized and abstract thought led them to lose this habit of visualizing if they ever had it.

After Galton's publication, two doctors published in quick succession the case of two of their patients. First, Professor Charcot reported the case of a patient who lost his ability to consciously evoke mental images following a stroke. This would also touch his unconscious imagery since he would not be able to dream with images as he used to. In addition, the patient suffered a loss of topographical memory and a prosopagnosia that is the inability to recognise the face of someone, even though the person is well known such as family or friends. When reporting the case, the term *amnésie visuelle* ("visual amnesia") would even be used to define this loss (***Bernard, 1883***). Nine years later, Doctor ***Wilbrand*** (***1892***) described a comparable case of a patient who had lost the ability of visual imagery (*Seelenblindheit* in the text) linked to a homonymous left hemianopia (partial blindness of half of the visual field) following an infarction in the occipito-temporal region. More than six decades later, ***Brain*** (***1954***) would state that this lost ability of mental imagery was induced by a loss of topographical memory and a severe prosopagnosia.

Once again, the topic disappeared from research interest for some decades until the early 1950s'. It is ***Critchley*** (***1953***) who linked these two cases and coined them under the term *Charcot-Wilbrand syndrome* (CWS) to denote the loss of ability to experience visual imagery in dreams following brain injury (***Botez et al., 1985***). Critchley mainly put the emphasis of CWS on defective re-visualisation or loss of visual images. The following year, ***Brain*** (***1954***) reported two more cases of CWS. The first patient had a car accident causing a depression fracture in the mid-frontal region and the second stumbled and struck his head, affecting the right fronto-vertical region.

Three decades later, ***Botez et al.*** (***1985***) reported the first case of a patient with a *pure innate* deficit of visual imagery without any other disorders.

More recently, experimental psychologist Bill Faw reported, among other things, his own case study by classifying himself as a "non-mental-imager" and how he learnt about it (***Faw, 2009***). In his paper, he commented on the denial of natural-non-imagers by his peers and contemporaries, despite the fact that the condition has been known for more than a century. He highlighted the fact that visual imagery is needed by most to understand their surroundings and that they can have difficulties understanding how others could function without it. Faw also briefly presented the case of a psychologist that lost her visual imaging ability following a car accident. This psychologist explained how, for six months, she struggled to understand people since she could not use her mind's eye to visualize what they meant anymore. Then, she managed to change her way

of processing what people told her and she moved from being an "imagery-decoder" to being an "auditory decoder". He also reports two large scales studies on 2500 and 750 people, that suggests between 2.1 - 2.7% of the population lack mental visual imagery. To find those numbers, Faw used derived versions of Galton's breakfast table questionnaire ((***Galton, 1880***)) and Bett's Questionnaire upon Mental Imagery (QMI) (***Betts 1909***; See Subjective measure for more information on the QMI).

Contemporaneously, ***Zeman et al.*** (***2010***) report the case of a patient (MX), who had lost the ability to produce visual images following coronary angioplasty. Even though MX lost the ability to produce visual images, he did not lose any other cognitive functions and became the first patient with a "pure" case of imagery generation disorder. Six years later, ***Zeman et al.*** (***2015***) suggested a new term for the absence of mental visual imagery (inspired by Aristotle), aphantasia from *a-* ("without") and *-phantasia* ("imagery"). The term "hyperphantasia" appeared shortly after to denote the other extreme, a photo-like visual imagery.

Recently there has been more detailed research addressing the prevalence of aphantasia in the general population (***Dance et al., 2022***). In this paper, they reported that 3.9% of the population has visual aphantasia, considering low and non-imagers, with 1 in 5 of them (0.8%) having a complete lack of visual imagery. To obtain those results, they used the Vividness of Visual Imagery Questionnaire (VVIQ) (***Marks 1973***; See Subjective measures for supplementary information on the VVIQ).

Nowadays, neuroimagery techniques have allowed us to better understand the functioning of visual mental imagery (***Pearson, 2019***).

Subjective measures

With the recognition of individual differences in visual imagery with Galton, research began exploring variability across other sensory modalities.

Hence, a number of self-report questionnaires have been developed and translated in to other languages in order to get more data on mental imagery. The most frequently used measure of visual mental imagery is the Vividness of Visual Imagery Questionnaire (VVIQ) (***Marks 1973***; see Appendix 1 Figure 2. This questionnaire includes 16 items that requires participants to produce mental images to answer on a 5-point scale. Other questionnaires measure a range of modalities such as the Vividness of Olfactory Imagery Questionnaire (VOIQ), to measure olfactory imagery (***Gilbert et al., 1998***), the Vividness of Movement Imagery Questionnaire (VMIQ), to measure kinaesthetic (or motor) imagery (***Roberts et al., 2008***) or the Clarity of Auditory Imagery Scale (CAIS), to measure auditory imagery (***Willander and Baraldi, 2010***). All these questionnaires are informative but are limited by their single modality approach.

Another questionnaire, developed in the early 20th century, takes into account seven sensory domains: the Questionnaire Upon Mental Imagery (QMI) (***Betts, 1909***). This questionnaire was updated, reducing the number of items from 150 to 35, with five items per modality and a 7-point rating scale (***Sheehan 1967***; See Appendix 1 Figure 1). This measure allows researchers to test participants auditory, cutaneous, emotional, gustatory, kinaesthetic, olfactory and visual mental imagery in just 10 minutes (versus 55 minutes for the original). The main advantage of the QMI over the measures presented above is the multi-modality approach. In a recent study at UGA (in preparation) preliminary analyses, conducted online on 893 participants, found that the VVIQ and the visual QMI were highly correlated using Spearman's Rho ($r_s = 0.93$, $p < .001$), just as the CAIS and the auditory QMI ($r_s = 0.88$, $p < .001$), confirming previous results from Pearson's team (***Dawes et al., 2020***).

Objective measures

More recently, researchers have been working on new ways of measuring visual mental imagery that do not include neuroimaging techniques. These have to remain objective measures and to be an equivalent of widely used questionnaires such as the VVIQ. These techniques implicate pupils.

Even though the pupil has a lot of functions, its main purpose remains to optimize vision. To do so, it constricts in response to brightness and to near fixation. On the other hand, it dilates in response to darkness but also in case of strong mental efforts and emotions (***Mathôt et al., 2018***). Now, assume that, if a mental image is a representation of a perception, the properties should be maintained in both imagery and perception. Since pupil diameter is, in part, linked to surrounding luminance, recalling a bright image should induce pupil constriction and recalling a dark image should induce pupil dilatation. (***Mathôt et al., 2017***)

Laeng and Sulutvedt (***2014***) created five experiments to explore this hypothesis. In experiment I, they showed triangles with one of eleven brightness levels, after a resting phase, participants were asked to revisualize the previously presented triangle. They found here that pupil size was correlated to remembered brightness through mental imagery. In experiment II, they compared active and passive imagery. In the first case, participants were asked to imagine the triangle previously observed, just as in experiment I. In the second case, participants observed triangles and then a grey background just as in experiment I. The only difference here is that they were not asked to revisualize triangles during the grey phase. The presence or absence of imagery instructions was the only difference there. They found there was no modification of pupil size during passive imagery. In experiment III, they reproduced previous results showing that pupil size could be a reliable index of the cognitive load. They presented different forms across two conditions, simple or complex, and then asked participants to imagine the stimuli. As expected, they found that the greater the cognitive load, the wider the pupil size became. In experiment IV, they controlled for potential experimental bias and potential voluntary control over pupil dilation. They asked participants to actively try to modify their pupil size. As expected, they found pupil dilation is not accessible to active control. In experiment V, they explored the extent to which differences in pupil dilation were triggered when using either short term memory (by recalling a form that had been seen previously) or long term memory. They presented different scenes in simple words that had dark or bright connotations. Then, asked participants to imagine those scenes. They found that pupil size was correlated with the imagined brightness or darkness of the scenario, consistent with perception. Ultimately, they devised a range of novel experiments and data analysis techniques for measuring visual mental imagery through pupil size. These techniques do not require neuroimagery nor questionnaires and can provide objective data of the strength of visual imagery.

Recently, Australian researchers used the same methodology as ***Laeng and Sulutvedt*** (***2014***) to measure visual aphantasia (***Kay et al., 2022***). They compared a group with typical visual imagery to a group of individuals with self-reported aphantasia. Kay & al. conducted two different experiments. The first one is a pupillary response task that was very similar to the first experiment of Laeng & Sulutvedt. The second task is a binocular rivalry task. In this experiment, participants were presented different patterns and color in each eye thanks to special glasses. They should report which pattern they perceived. Two conditions were opposed. One with no priming that should give a 50% chance to obtain one response or the other. The other one with a priming where participants have to think of one of the two colours before they are presented with the patterns (***Pearson et al., 2008***; ***Pearson, 2014***). Also, they used the VVIQ to compare the objective results from their experiments with widely used subjective ones.

Goals and hypothesis of our study

In this paper, we aim to reproduce ***Kay et al.*** (***2022***) findings. We mainly focus on the pupillary response task during visual mental imagery with two experiments. The first experiment is very similar to the original experiment proposed by Kay. The second experiment mimics the first one but perception phase and imagery phase are reversed (see Procedure). After each trial of the two experiments, participants gave a score to indicate the vividness of their imagery. We will compare those results with the (visual) QMI score. We choose to use the QMI instead of the VVIQ because

further concurrent studies are being conducted within this team that explore alternative objective measures of auditory-verbal mental imagery (inner speech). Using the multi-modality of the QMI allows us to compare our results across studies using a consistent measure. Moreover, it has been shown that the VVIQ and the visual QMI are highly correlated (see Subjective measures).

We also reproduce the Pupil Cycle Time experiment from ***Lamirel et al.*** (***2018***) (inspired by ***Miller and Thompson 1978***). Since the PCT is a really short experiment, we decided to extend the protocol with additional conditions. Nevertheless, we did not have enough time to analyse this data. We reported our method in the "Methods and Materials" section, but we did not give it a part in the "Result" section.

As found in previous research, we expect the pupil to dilate when participants are imagining dark stimuli and a constriction of the pupil when participants are imaging bright stimuli. Since we ask participants to report the vividness of their visual imagery both at the end of each trial and with the QMI, we will obtain one mean rating per participants and per pupillometry experiment plus another one from the QMI. We expect a positive linear correlation between the visual QMI and both mean ratings. Moreover, we expect a direct link between self-reported visual imagery and the difference in pupil size between the two conditions (bright and dark). In other words, we expect participants with higher QMI scores (30 and above) to have the biggest differences in pupil size while imagining bright stimuli and dark stimuli and participants with lower QMI scores (15 and below) to have none. Moreover, we expect participants to have greater pupil differences when they report having more vivid imagery during trials.

Results

This section is segregated in three sub-parts. First, we analyse our subjective data with the Questionnaire upon Mental Imagery results. This allows us to have a better understanding of our population sample. Then, we analyse our objective data, starting with the perception first (PF) experiment and then the imagery first (IF) experiment.

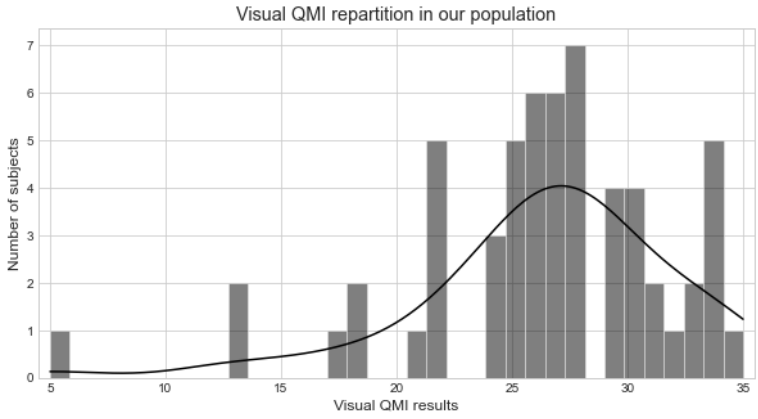
Questionnaire upon Mental Imagery

Questionnaire upon Mental Imagery (QMI) analysis were driven on the whole datasets (see Exclusion criteria). As expected in any group when it comes to measuring mental visual imagery, participants were not normally distributed along the visual QMI results. In fact, when looking at our results (see Figure 1a), we obtained more of a left-skewed distribution like ***Dance et al.*** (***2022***) had. But since just looking at a graph is not enough to get a full understanding of data, we processed some statistical analysis. We computed a skewness test on our data and obtained a score of -1.25 (Dance et al. skewness score: -0.65). Since this is a negative value, it confirms that our dataset is left skewed. We also computed the kurtosis that gave us 5.58 (Dance et al. kurtosis: 3.81). Since it is greater than 3, it means that there are more values in the tail than it would be expected in a normal distribution. Then, in order to be sure of our previous results, we computed the Jarque-Bera normality test for which the null hypothesis is that the dataset has a skewness and kurtosis that matches a normal distribution. We obtained a JB of 31.30 with a p-value of 1.59e-07 (Dance and colleagues results: JB = 97.751, p-value < 2.2e-16). Since the p-value is less than α = 0.001, we can reject the null hypothesis and confirm the skewness of our dataset. Finally, our dataset had a mean value of 26.34 for a standard deviation of 5.78 and a median of 27. Since Dance et al. used the VVIQ with a scale going up to 80 and we used the visual QMI with a scale going up to 35, we had to normalise both datasets in order to be able to compare both of them. After normalisation, we obtained a mean value of 0.75, a standard deviation of 0.16 and a median = 0.77 (Dance et al. results: μ = 0.72, sd = 0.15 and median = 0.725).

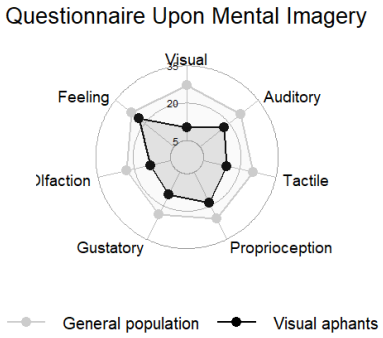
Our population counted 1 participant at floor with a score of 10/35 meaning that the participant always checked the answer *Je peux y penser mais je ne vois pas d'image* ("I think of it, but do not have an image before me") and 1 participant scoring at ceiling with a score of 35/35 meaning

	Visual	Auditory	Tactile	Proprioception	Gustatory	Olfaction	Feeling	Total
Mean	26,34	25,36	24,86	25,57	23,57	22,57	26,80	175,07
Standard deviation	5,78	5,73	5,52	5,26	6,55	6,99	5,19	31,21
Median	27	25	25	26	24	23	27	171
Floor	1	0	0	0	1	0	0	0
Ceiling	1	2	2	2	2	1	1	0
Moderate aphantasics	3	3	2	2	5	6	1	2
Moderate hyperphantasics	11	10	5	9	9	9	15	9

Table 1. Results of the Questionnaire upon Mental Imagery (N = 58 participants). Added to mean, median and standard deviation are the number of persons that scored at floor (always checked 1/7), the number of persons that scored at ceiling (always checked 7/7), the number of person that had moderate aphantasia (checked 3/7 or below) and the number of person that had moderate hyperphantasia (checked 6/7 or above)



(a)



(b)

Figure 1. Results obtained for the Questionnaire upon Mental Imagery (N = 58 participants). (1a) Distribution of visual QMI results of our population sample. (1b) Radar plot of the average QMI response across all modalities in the general population (N = 55 participants) and in our sub-population with moderate to complete visual aphantasia (N = 3)

Figure 1–Figure supplement 1. Comparison in self reported visual imagery

that the participant always checked the answer *Très clairement et nettement, comme dans la réalité* ("Very vivid and as clear as reality") (see Table 1 for more details).

Objective measures

Both of our experiments have five phases (based on ***Laeng and Sulutvedt 2014*** and ***Kay et al. 2022***; see Appendix 2-Figure 1). The baseline, that is at the beginning of each trial, allows us to measure the size of the participant pupil across the experiment. It is used to normalise pupil size across trials and participants. This measure is essential to compare both our participants against each other and each of their trials among them. The perception phase contains a triangle from one of two conditions, "black" or "white" (see Appendix 2-Figure 4 for triangles used). The resting phase is for the participants' pupil to come back to a standardised size. The imagery phase where the participants have to visualise the triangle they have previously seen. At the end of each trial, participants had to rate the vividness of their imagery during the last part.

The perception first (PF) experiment has the each phase in the order that is described above (see Appendix 2-Figure 2). The imagery first (IF) experiment is a modified version of the PF experiment where imagery and perception are inverted (see Appendix 2-Figure 3). Because the imagery phase

comes before the perception phase, participants are informed of the luminance they should imagine before the baseline in an auditory form through speakers.

In order to plot the evolution of pupil size across time (see Figure 2a and Figure 3a), data were pre-processed. After reconstructing pupil size during blinks (see Data analysis), results were normalised by dividing pupil size of each trial by the mean pupil size of the last 0.5 s of the baseline. Then, two average pupil sizes over time were computed per participant: one for the black condition and one for the white condition.

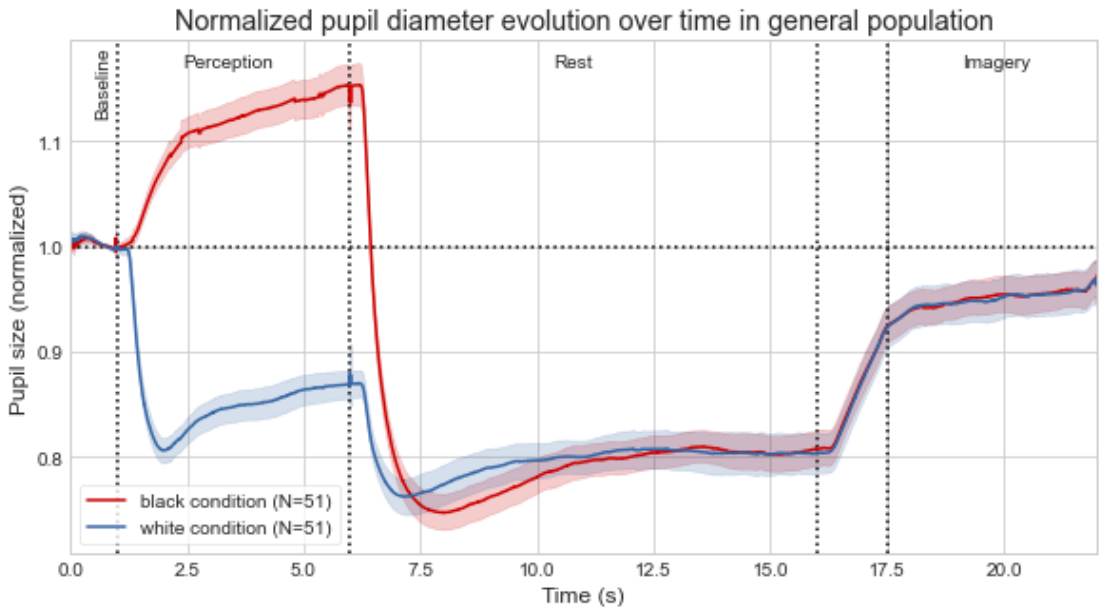
Perception first pupillometry experiment

In Figure 2, we can observe that baselines (between 0 - 1 s) are all at the same level. This is absolutely normal since we have divided each trial data by the mean value of its baseline (See Data analysis). Then, the perception phase (between 1 - 6 s) starts. Nothing happens for the first 0.2 s. After this short latency, the pupil will strongly constrict if the stimulus is white and will dilate if the stimulus is black. Constriction can be up to twice as fast as dilatation. (***Theodossiadis et al., 2012***). We observe a bouncing effect of the pupil in the white condition around 2 s called *pupil escape* (***Krenz and Stark, 1985***). Then, the pupil adjusts optimally to the new luminance by slowly dilating until the end of the phase. When the resting phase (between 6 - 16 s) starts, there is a new latency period of 0.2 s. Then, we have a pupil escape effect once again. This time, pupil diameter stays low. By the end of the phase, there is no difference in pupil diameter between conditions. Finally, the imagery phase (between 16 - 22 s) starts and, after the latency period, pupil dilates for 1.5 s. Afterwards, we observe that there is no difference in pupil size between black and white condition.

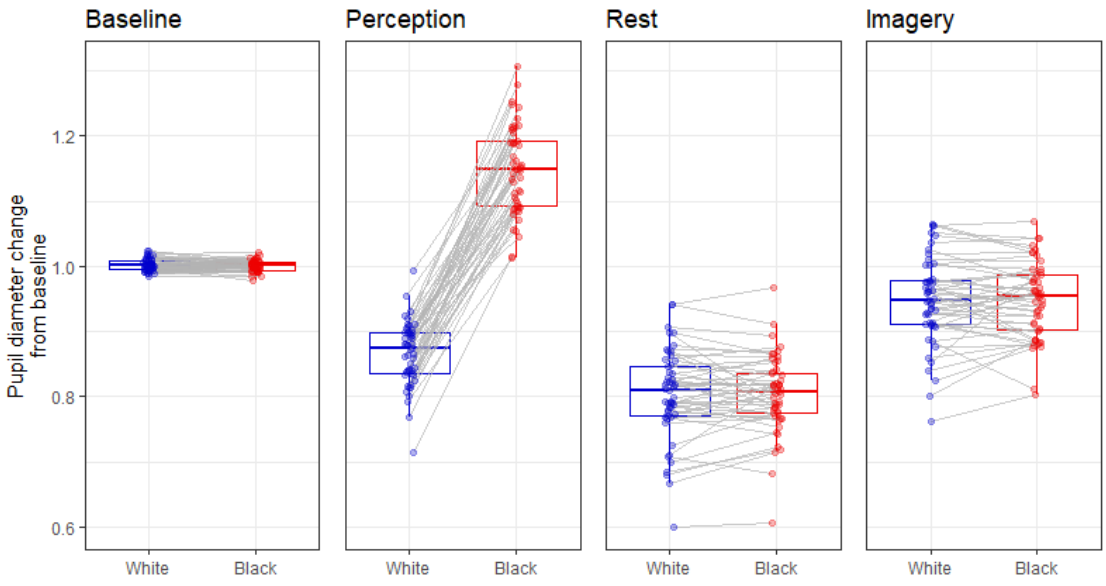
In order to have a deeper look at the data, we show boxplots of the pupil size of our participants during each phase for the black and the white conditions (see Figure 2b). Because pupil size evolves differently depending on the phase, we draw the boxplots on shorter periods. For the baseline, we use the last 0.5 s. For the perception phase, we use the last 1 s that corresponds to the optimal pupil size of the phase. For the resting phase, we take the last 1 s to be sure that pupils of our participants have enough time to reach the same diameter. For the imagery phase, we take the 4 s between the end of the first auditory cue and the beginning of the second auditory cue. There is a significant difference in pupil size during perception depending on conditions "black" and "white" ($t = 32.84$, $p < 0.001$). As expected, pupils of participants come back to the same size depending on conditions "black" and "white" by the end of the resting phase ($t = 0.10$, $p = 0.919$). Nevertheless, pupil size does not differ during the imagery phase depending on conditions "black" and "white" ($t = 0.07$, $p = 0.943$). This absence of effect is even retrieved in participants with visual QMI scores of 30 or more (see Figure 2-Figure supplement 1). Results are also analysed depending on mean rating. Participants rating across the experiment are averaged to keep only their mean rating. Then, they are segregated in two subgroups. One group with low mean rating with an average score of 2 or less (see Figure 2-Figure supplement 2) and one group with high mean rating with an average score of 3 or more (see Figure 2-Figure supplement 3). We find no significant effect in one case or the other.

Imagery first pupillometry experiment

In Figure 3, we can observe that baseline (between 0 - 1.6 s) are all at the same level. Just as in the previous experiment, it is absolutely normal since we have divided each trial data by the mean value of its baseline. Then, the imagery phase (between 1.6 - 7.6 s) starts. Nothing happens for the first 0.2 s. After the latency period, pupil slowly constricts. As a reminder, the edges of the triangle presented during this phase were slightly darker than the background but also slightly lighter than the fixation cross and thinner. From the beginning to the end of this phase, there is no sign of difference in pupil size depending on condition. During the first resting phase, there is a small dilatation quickly followed by a constriction. This phenomenon is due to the fact that the fixation



(a)



(b)

Figure 2. Results obtained for the perception first experiment (N = 51 participants). (2a) Mean pupil size waveforms of our population sample during the experiment. The lines represent the mean value and the shadows represent the standard error. (2b) Average pupil size during each phase depending on the condition. Each point represents a participant in a condition. A grey line links the mean pupil size of a participant in one condition to the mean pupil size of the same participant in the other condition.

Figure 2–Figure supplement 1. Pupil evolution across time for the perception first experiment for individuals with low QMI results

Figure 2–Figure supplement 2. Pupil evolution across time for the perception first experiment for individuals with low QMI results

Figure 2–Figure supplement 3. Pupil evolution across time for the perception first experiment for individuals with low QMI results

cross replaces the edges of the triangle and is slightly darker. Pupil size does not differ from one condition to the other by. The perception phase has the same profile as the perception phase of the previous experiment. There is a latency period of 0.2 s followed by a strong constriction in the white condition and a dilatation on the perception phase. Then, the pupil slowly dilates back to its optimal diameter. Finally, after the usual latency, the pupil comes back to a same level of aperture during the second resting phase.

In order to have a deeper look at the data, we show boxplots of the pupil size of our participants during each phase (see Figure 3b). As for the previous experiment, we draw the boxplots on shorter periods. For the baseline, we use the last 0.5 s. For the imagery phase, we take the 5 s between the end of the first auditory cue and the beginning of the second one. For the perception phase, we use the last 1 s that corresponds to the optimal pupil size of the phase. For both resting phases, we take the last 1 s to be sure that pupils of our participants have enough time to reach the same diameter.

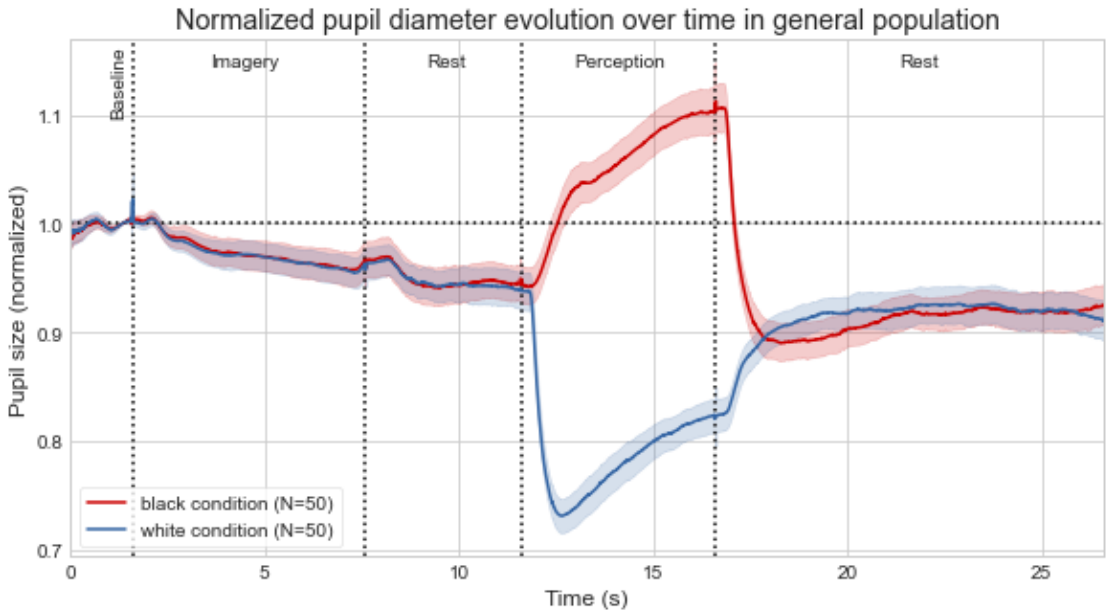
As in the previous experiment, there was a difference in pupil size during perception of the triangles from each condition (see Figure 3b; t = -0.13, p = 0.55). Also, pupils of participants came back to the same size depending on conditions "black" and "white" by the end of the first resting phase (t = 0.97, p = 0.17) and the second resting phase (t = 0.69, p = 0.25). When it came to imagery, we also obtained the same results than in the PF experiment with no significant difference depending on conditions (t = 0.25, p = 0.40). This absence of effect is even retrieved in participants with visual QMI scores of 30 or more (see Figure 3-Figure supplement 1). Results were also analysed depending on mean rating. Participants rating across the experiment were averaged to keep only their mean rating. Then, they were segregated in two subgroups. One group with low mean rating with an average score of 2 or less (see Figure 3-Figure supplement 2) and one group with high mean rating with an average score of 3 or more (see Figure 3-Figure supplement 3). We found no significant effect in one case or the other.

Discussion

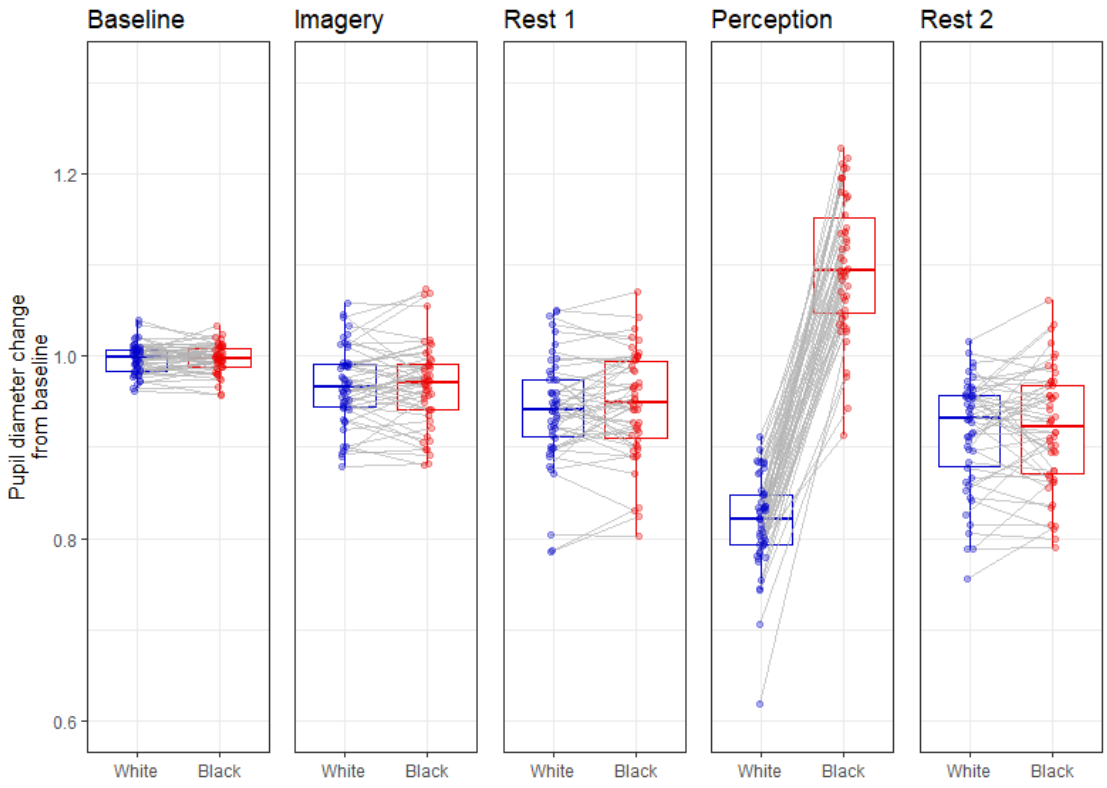
Our experiments do not confirm previous results that the pupil dilates when participants are visualising dark stimuli and constricts when they were visualizing bright stimuli (***Kay et al., 2022***). Since we could not find a link between imagery and pupillary response, we do not observe a direct link between self-reported visual imagery and pupil size depending on conditions. In fact, most of our subjects reported having difficulties visualising triangles when we asked them after the experiments. This can be confirmed by the fact that visual QMI and mean rating during PF experiment were not correlated (r = 0.33, p = 0.10). During the IF experiments, reported mean rating was correlated to visual QMI (r = 0.25, p = 0.01). Our participants also reported having more ease to imagine triangles when they had seen it before (with a mean rating μ = 2.66, sd = 0.07), during the PF experiment than to fill in the triangle with color at the beginning of the trial (with a mean rating μ = 2.40, sd = 0.10), during the IF experiment. In other words, our subjects had more ease to recall an image they had seen before than to create a luminance level before their mind eye. This self report from our subject is also seen in rating data (see Figure 1-Figure supplement 1).

There were a couple of discrepancies between ***Kay et al.*** (***2022***) and our experiment.

First, we did not use the same eye tracker than Kay et al. In our study, we used the Eyelink 1000 device. According to its user manual user manual, it has a noise level of 0.2% of the diameter, corresponding to a resolution of 0.01 mm for a pupil of 5 mm. Since Kay et al. found an average difference of 0.21 mm in pupil size during imagery between their dark and bright conditions, it should not be a problem. On their side, Kay et al. used head mounted eye-tracking glasses from the Pupil Labs that were originally created to provide accurate data on gaze in more natural situations since it allows the subject to move his or her head (***Kassner et al., 2014***). This technology is not known to be more accurate on pupil size. In 2019, researchers compared Eyelink 1000 and Pupil Labs glasses and found differences between the two at the individual subject level but not on the



(a)



(b)

Figure 3. Results obtained for the imagery first experiment (N = 50 participants). (3a) Mean pupil size waveforms of our population sample during the experiment. The lines represent the mean value and the shadows represent the standard error. (3b) Average difference of pupil size during each phase depending on the condition. Each point represents a participant in a condition. A grey line links the mean pupil size of a participant in one condition to the mean pupil size of the same participant in the other condition.

Figure 3–Figure supplement 1. Pupil evolution across time for the imagery first experiment for individuals with high QMI results

Figure 3–Figure supplement 2. Pupil evolution across time for the imagery first experiment for individuals with low mean rating results

Figure 3–Figure supplement 3. Pupil evolution across time for the imagery first experiment for individuals with high mean rating results

group level ***Ehinger et al.*** (***2019***). Two important points have to be flagged from on Ehinger et al. experiment. First, they had a small sample size (15 participants) that might not be enough to draw conclusions on a larger scale. Second they did not measure the true pupil size during the experiments. Overall, their findings suggest that Kay's results should not be impacted. Except for their within-subjects findings that suggest a link between pupil size and rating of individuals on their imagery during the trials. However, it had to be recalled in case future researchers would like to use one technology or another at an individual subject scale.

Methodologically speaking, here is a list of the differences between Kay and colleagues experiment and our perception first experiment that was supposed to mimic it. We used a dark grey fixation cross while they were using a white one. We used a uniform dynamic mask of 10 s for the resting phase where they used a dark screen of 8 s for their resting phase. We had two conditions with one color per condition (either black or white). That means that we showed two times each triangle to our participants. Kay et al. had also two conditions but they had two brightness levels per condition: black and dark grey for their "dark" condition and white and light grey for their "bright" condition. That means that their participants saw each stimulus only once, minimizing boredom. We asked our participants to rate the vividness of their mental image with the mouse where Kay et al. used the keyboard. Finally, Kay and colleagues used another condition. Their stimuli were either "Set-Size-One" or "Set-Size-Four" to test for cognitive load.

Interestingly, our sample had a similar distribution than ***Dance et al.*** (***2022***) when they tested more than a thousand persons. This fact implicates that our sample was representative of the general population and reinforces previous findings on the prevalence of aphantasia in population.

One limitation of our study is that it was conducted on psychology students. Even it has been shown that there were representative of the general population on their subjective reports of their own visual mental imagery, there might be other parameters that were not taken into account. Also, we did not check for a lack of active engagement. This could have been done by including tasks that requires a higher cognitive load. That is what ***Kay et al.*** (***2022***) did in their study by including stimuli with 4 triangles instead of only 1 as we did.

The main finding of this paper is the fact that pupil size does not only depend on average brightness during perception. In fact, while passively looking at the resting phase during the PF experiment, participants pupil were significantly more contracted than during the baseline or the imagery phase (see Figure 2b). Even though the average luminance of the screen was the same for all three phases, their variability was different. During baseline and imagery, there was a uniform grey background. During rest, there was a uniform dynamic mask (see Appendix 2-Figure 2) that refreshed with every frame (that is every 0.013 s). Nevertheless, both backgrounds had the same mean luminance value overall. That means that pupil size does not solely depends on average brightness but also on brightness variability.

To conclude, we did not find that visual mental imagery of brightness had an impact on pupil size in both short-term memory (with the PF experiment) and long-term memory (with the IF experiment) conditions. It may be interesting for future research to further investigate the differences in visual imagery between short-term memory and "pure" brightness imagery when it comes to pupillary response. Even though we did not replicate previous results in this paper, ***Laeng and Sulutvedt*** (***2014***) and ***Kay et al.*** (***2022***) findings are not to throw away and a proper replication should be conducted. If it appears that tiny changes in our methodology can lead to a replication of previous results, it will be possible to conduct the experiment on larger group of individuals with aphantasia. Our team accumulated contacts from hundreds of people without visual imagery, French and English, that would be interested in participating in further research. Identifying unbiased neurophysiological measures of visual imagery like this one could help us to better understand it and help to the development of theories that are more favourable to the mind's eye.

Methods and Materials

Population

Fifty-eight naive participants with a mean age of 21 (range 18-30, 57 women) participated in the experiments. All participants were psychology students and were given course credits for their participation in our study. All participants had normal or corrected to normal vision. Due to exclusion criteria, all of the participants could not be kept in our data analysis. Ultimately, 51 participants were included in the PF experiment and 50 participants were included in the IF experiment. All participants were kept for the QMI (See Exclusion criteria for more details).

In order to observe a difference in average between the conditions, a Cohen's d have been calculated using G*Power (***Faul et al. 2007***; version 3.1.9.7). It has been calculated that a sample size of 54 participants was required in order to find a difference between two dependent means with a two tails test (g*Power effect size dz = 0.5, α = 0.05, β = 0.95). Those parameters are the ones used by ***Kay et al.*** (***2022***) in his experiment.

Material

All visual stimuli were displayed by a Dell desktop (DELL Precision 5820 Tower) linked to a Dell cathodic monitor (Dell M993s) of size 360 * 275 cm, with 75 Hz refresh rate and a 1280 * 1024 resolution. Contrast was set to 100% and brightness to 50%. Screen luminance was measured using a Spyder Elite 5 photometer and its gamma trend curve was fitted in MATLAB (version R2017b). With Y being the measured luminance of the screen and X being the grey RGB value from 0 to 255 (see Equation 1). The R^2 is the correlation coefficient that mirrors the goodness of fit of a model on the data. To be more precise, this statistical measure represents the proportion of the dependent variable (here Y, the measured luminance of the screen) that can be explained by the independent variable (here Y, the RGB value that is presented on screen). That measure ranges from 0-1 and so an R^2 of 0.999 represents a really good fit of the data from the equation.

$$Y(X) = 2.652 + 0.003 * X^{1.863} \quad (1)$$
$$R^2 = 0.999$$

Since pupil is known to change in size depending on surrounding brightness, experiments were conducted in a darkened room to remove any possible fluctuation in lighting. Participants had to place their head in a head rest situated at 65 cm from the screen to limit head movements.

Eye gaze and pupil diameter of the right eye were recorded at 1000 Hz using a video-based eye tracker EyeLink 1000 Plus (SR Research, Mississauga, ON, Canada), using infrared lights and with a 35mm monocular lens.

Participants were given a link to fill in the QMI (https://enquetes-screen.msh-alpes.fr/index.php/927978?lang=en; see Appendix 1-Figure1).

Pupillometry imagery experiments were created using PsychoPy Builder (v2021.2.3), modified with Python (v3.9) and run with PsychoPy Runner (v2021.2.3). The PCT experiment was created in MATLAB (version R2017b). The QMI was created on LimeSurvey.

Stimuli

For both pupillometry imagery experiments, 8 achromatic triangles were created. Triangles were evenly divided in two groups and were either "black" or "white". Black triangles had a luminance of 1.6 cd/m². White triangles had a luminance of 86.28 cd/m². Triangles had four possible orientations, either 0°, 90°, 180° or 270°. All triangles were equilateral with a size of 12.5 degrees of visual angle (or 14.25 cm sides on screen in our case). During experiments, every triangle was randomly presented twice.

For the perception first (PF) experiment, a dynamic mask, changing with each frame, was created with pixel luminance uniformly ranging from 1.6 cd/m² to 86.28 cd/m². This mask was displayed on the whole screen during the resting phase (see Appendix 2-Figure 2).

For the imagery first (IF) experiment, triangle contours had the same size and possible orientations than triangles described above. Contours were 1 px large and had a luminance of 21.65 cd/m². The terms *noir* ("black") and *blanc* "white" were recorded by the experimenter via Audacity (v3.1.3).

For both pupillometry imagery experiments, a neutral grey background with a luminance of 25.46 cd/m² was used. The fixation cross had a luminance of 18.22 cd/m², a size of 0.275 cm and a thickness of 1 px. All stimuli described above were created using PsychoPy build-in functions.

For the Pupil Cycle Time experiment, a blue fixation cross of 1° of visual angle and a RGB value of [0,0,200] was used. The white disk had a diameter of 270 cm. At first, its RGB value was 127.5. Its luminance changed every frame depending on the pupil diameter of the participant (see Equation 2). With Maximum_pupil_size being recorded at the beginning of the experiment where the participant should just look at a black screen for 12 seconds. All of the PCT experiment and its stimuli were created in MATLAB.

$$Disk\_colour = Current\_pupil\_size / Maximum\_pupil\_size * 255 \quad (2)$$

Procedure

Three different experiments were conducted. The first experiment was a pupillometry imagery experiment, inspired by ***Laeng and Sulutvedt*** (***2014***) and ***Kay et al.*** (***2022***), that we named perception first (PF) experiment. The second experiment was a reversed version of the previous one with imagery and perception being inverted. This one was named imagery first (IF) experiment. The third experiment was an exploratory version of the Pupil Cycle Time (PCT) (***Lamirel et al., 2018***). The order of experiments 1 and 2 was counterbalanced. Experiment three was always run after the two other ones. Before each experiment, participants were reminded to fix the centre of the screen and to minimise blinking during trials. After all three experiments, participants had to fill in the Questionnaire upon Mental Imagery (QMI). Instructions of experiments one and two were verbally explained with the help of a PowerPoint (Office v18.2110.13110.0) presentation. The concept of after image was explained and a scale of vividness was presented for participants to have a reliable idea of how to rate their mental visual imagery. Instructions were also written on the screen before each experiments start. Instructions of the PCT were verbally explained and written on the screen before each trial.

Perception first experiment timeline: Each trial began with a dark-grey fixation cross at the center of a neutral grey screen (baseline) for 1 s. A triangle from one of two conditions (black or white; see Appendix 2-Figure 4). was then presented at the center of the neutral grey screen for 5 s. During this time, participants were instructed to focus on the triangle to memorise its size, orientation and luminance. Then, a fixation cross was displayed on a uniform dynamic mask for 10 s to maximally reduce after images. The baseline screen was presented again for 6 s. The beginning and the end of this phase were both marked by an auditory beep (440 Hz for 500 ms). Between those auditory signals, participants were asked to imagine or to visualise the triangles they were presented earlier in the trial. Finally, participants were asked to rate the vividness of their imagery, between the auditory signals, on a scale of 1-4, 1 being *Pas du tout vive - Aucune forme n'est apparue dans l'imagerie* ("Not vivid at all – No shape appeared in imagery") and 4 being *Très vive - presque comme si elle avait été vue* ("Very vivid – Almost like seeing it"). No labels were used for points 2 and 3. Answers had to be checked using the mouse and answering lead directly to the next trial. Two training trials were used to be sure that participants understood the experiment correctly.

Image first experiment timeline: Each trial began with a dark-grey fixation cross at the center of a neutral grey screen (baseline) for 1.6 s. At the beginning of this phase, the term *noir* ("black") or *blanc* "white" was aurally heard by the participant. The outer edges of a triangle was then presented on the neutral grey screen for 6 s. Contours were slightly darker than the background. The

beginning and the end of this phase were both marked by an auditory beep (440 Hz for 500 ms). Between those auditory signals, participants were asked to imagine filling the triangle with the colour they heard at the beginning of the trial. The baseline screen was presented again for 4 s. A triangle, corresponding to the one they were supposed to visualise during the trial was displayed on a neutral grey screen for 5 s. The baseline screen was then presented again for 10 s. Finally, participants were asked to rate the vividness of their imagery, between the auditory signals, on a scale of 1-4, 1 being *Pas du tout vive - Aucune forme n'est apparue dans l'imagerie* ("Not vivid at all – No shape appeared in imagery") and 4 being *Très vive - presque comme si elle avait été vue* ("Very vivid – Almost like seeing it"). No labels were used for points 2 and 3. Answers had to be checked using the mouse and answering lead directly to the next trial. Two training trials were used to be sure that participants understood the experiment correctly.

Pupil Cycle Time: In the first trial, a blue fixation cross was presented on a black screen (baseline) for 10 s. In the second trial, the baseline screen was presented for 2 s. Then, a white disk was presented behind the fixation cross for 30 s. The luminance of the disc varies inversely with the diameter of the pupil. The third and fourth trial were similar to the second one in the method but while participants were looking at the center of the screen, they were asked to imagine that the screen was either of a deep black (*d'un noir profond*) or extremely bright (*extrêmement brillant*). After trial three and four, participants were asked to rate the vividness of their imagery on a scale of 1-4, 1 being *Pas du tout vive - Aucune forme n'est apparue dans l'imagerie* ("Not vivid at all – No shape appeared in imagery") and 4 being *Très vive - presque comme si elle avait été vue* ("Very vivid – Almost like seeing it"). No labels were used for points 2 and 3. Answers had to be checked using the mouse. Participants had to tap the space bar whenever they were ready to start a new trial. The order of the last two trials was randomly determined. There was no training trial in this experiment.

Exclusion criteria

We did not want our participants to be on any kind of drugs and so two were removed from being on antidepressant because we did not know the effects of these on their pupil size. One more participant was removed because the participant eyes were jumping from left to right when the participant had to focus on the center of the screen.

Three participants were removed from analysis because they had a score of less than 15/35 (corresponding to *Vague et peu clair* ("Vague and dim")) for the visual QMI. This threshold is based on the thresholds usually used for the VVIQ, main questionnaire used to measure one lack of visual imagery (***Zeman et al.*** (***2015***); ***Fulford et al.*** (***2018***); ***Keogh et al.*** (***2021***); ***Kay et al.*** (***2022***); ***Dance et al.*** (***2022***), that considers that people have visual aphantasia or moderate visual aphantasia when they checked "Vague and dim" or less. Overall, six participants were removed from both experiments.

One participant's data were not recorded.

One participant was removed from the IF experiment because the fire alarm was triggered while the participant was going through the experiment and the participant could not finish it. Nevertheless, this participant was kept for the PF experiment.

Out of 58 participants, 51 participants were included in the PF experiment and 50 participants were included in the IF experiment. (See Appendix 1-Figure 5). We kept all of our participants in the QMI analysis because we just wanted to define our population with it.

Data analysis

To preprocess pupillometric data, we mainly used (***Mathôt, 2018***) guidelines and codes. In order to estimate pupil size during blinks, we used cubic spline interpolation (***Mathôt, 2013***). Since pupil data contain a lot of artifacts (***Mathôt et al., 2018***; ***Kret and Sjak-Shie, 2019***), we used a moving Hanning window of 51 ms as prescribed by ***Mathôt et al.*** (***2017***). Data were then normalised in order to be comparable between each trial and between each subject. To do so, we divided each trial by its baseline (i.e. the last 500 ms before starting the perception phase in the PF experiment

and the last 500ms before starting the imagery phase in the IF experiment). Repeated baselines prevent effects of fatigue on pupil size across the experiment ***Kret and Sjak-Shie*** (***2019***); ***Mathôt*** (***2018***). All of this preprocessing was run on Python (v3.9).

Paired sample t-tests for PF and IF experiments were run with Jamovi (v2.3.9). QMI data were analysed with RStudio (v2021.09.1). Correlation between visual QMI and mean rating was performed in Python.

Pupil cycle time was explored in MATLAB. At the moment, we have not finished to analyse the data and so we decided not to discuss them in this paper.

Acknowledgments

I would like to express my gratitude to my supervisors CHAUVIN Alan and GUYADER Nathalie for giving me the opportunity to work on this project and helping me to realise it, from implementation to final review. A special thanks to my colleagues in the Aphantasia team with HUSON Nicole who reviewed my paper, PESCI Téo and VANBUCKHAVE Claire who were always present for answering any question I could have, and FABER Léa and LOEVENBRUCK Hélène who followed the project from distance with a critical view.

I would also like to thank my friends, and fellow trainees, wholeheartedly for their support and kindness. Finally, I would like to thank my sister for her review, my girlfriend for her patience and support, and, last but not least, my flatmates for their pasta supplies throughout the semester.

References

Bernard DAF. Un cas de suppression brusque et isolée de la vision mentale des signes et des objets : formes et couleurs. Clinique des Maladies Nerveuses. 1883; p. 22.

Betts G Herbert. The distribution and functions of mental imagery. New York: Comlumbia University; 1909.

Botez MI, Olivier M, Vézina JL, Botez T, Kaufman B. Defective Revisualization: Dissociation Between Cognitive and Imagistic Thought Case Report and Short Review of the Literature. Cortex. 1985 Sep; 21(3):375–389. https://linkinghub.elsevier.com/retrieve/pii/S0010945285800034, doi: 10.1016/S0010-9452(85)80003-4.

Brain R. Loss of Visualization. Proceedings of the Royal Society of Medicine. 1954 Apr; 47(4):288–290. http://journals.sagepub.com/doi/10.1177/003591575404700410, doi: 10.1177/003591575404700410.

Critchley M. The parietal lobes. London: Edward Arnold Ltd; 1953.

Dance CJ, Ipser A, Simner J. The prevalence of aphantasia (imagery weakness) in the general population. Consciousness and Cognition. 2022 Jan; 97:103243. https://linkinghub.elsevier.com/retrieve/pii/S1053810021001690, doi: 10.1016/j.concog.2021.103243.

Dawes AJ, Keogh R, Andrillon T, Pearson J. A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. Scientific Reports. 2020 Dec; 10(1):10022. http://www.nature.com/articles/s41598-020-65705-7, doi: 10.1038/s41598-020-65705-7.

Ehinger BV, Groß K, Ibs I, König P. A new comprehensive eye-tracking test battery concurrently evaluating the Pupil Labs glasses and the EyeLink 1000. PeerJ. 2019 Jul; 7:e7086. https://peerj.com/articles/7086, doi: 10.7717/peerj.7086.

Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods. 2007 May; 39(2):175–191. http://link.springer.com/10.3758/BF03193146, doi: 10.3758/BF03193146.

Faw B. Conflicting Intuitions May Be Based On Differing Abilities. Journal of Consciousness Studies. 2009 Jan; 16(4):25.

Fulford J, Milton F, Salas D, Smith A, Simler A, Winlove C, Zeman A. The neural correlates of visual imagery vividness – An fMRI study and literature review. Cortex. 2018 Aug; 105:26–40. https://linkinghub.elsevier.com/retrieve/pii/S0010945217303209, doi: 10.1016/j.cortex.2017.09.014.

Galton F. Mind a quarterly review of psychology and philosophy. Mind. 1880 Jan; os-V(19):301–318. doi: 10.1093/mind/os-V.19.301.

Gilbert AN, Crouch M, Kemp SE. Olfactory and visual mental imagery. Journal of Mental Imagery. 1998; 22(3-4):137–146. https://psycnet.apa.org/record/1998-03238-006.

Gregoric P. Aristotle on the common sense. Oxford Aristotle studies, Oxford ; New York: Oxford University Press; 2007. OCLC: ocn176911656.

Kassner M, Patera W, Bulling A. Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* Seattle Washington: ACM; 2014. p. 1151–1160. https://dl.acm.org/doi/10.1145/2638728.2641695, doi: 10.1145/2638728.2641695.

Kay L, Keogh R, Andrillon T, Pearson J. The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength. eLife. 2022 Mar; 11:e72484. https://elifesciences.org/articles/72484, doi: 10.7554/eLife.72484.

Keogh R, Pearson J, Zeman A. Aphantasia: The science of visual imagery extremes. In: *Handbook of Clinical Neurology*, vol. 178 Elsevier; 2021.p. 277–296. https://linkinghub.elsevier.com/retrieve/pii/B978012821377300012X, doi: 10.1016/B978-0-12-821377-3.00012-X.

Krenz WC, Stark L. Systems model for pupil size effect: II. Feedback model. Biological Cybernetics. 1985 Apr; 51(6):391–397. http://link.springer.com/10.1007/BF00350779, doi: 10.1007/BF00350779.

Kret ME, Sjak-Shie EE. Preprocessing pupil size data: Guidelines and code. Behavior Research Methods. 2019 Jun; 51(3):1336–1342. http://link.springer.com/10.3758/s13428-018-1075-y, doi: 10.3758/s13428-018-1075-y.

Laeng B, Sulutvedt U. The Eye Pupil Adjusts to Imaginary Light. Psychological Science. 2014 Jan; 25(1):188–197. http://journals.sagepub.com/doi/10.1177/0956797613503556, doi: 10.1177/0956797613503556.

Lamirel C, Ajasse S, Moulignier A, Salomon L, Deschamps R, Gueguen A, Vignal C, Cochereau I, Lorenceau J. A novel method of inducing endogenous pupil oscillations to detect patients with unilateral optic neuritis. PLOS ONE. 2018 Aug; 13(8):e0201730. https://dx.plos.org/10.1371/journal.pone.0201730, doi: 10.1371/journal.pone.0201730.

Marks DF. Visual imagery differences in the recall of pictures. British Journal of Psychology. 1973 Feb; 64(1):17–24. https://onlinelibrary.wiley.com/doi/10.1111/j.2044-8295.1973.tb01322.x, doi: 10.1111/j.2044-8295.1973.tb01322.x.

Mathôt S. A simple way to reconstruct pupil size during eye blinks. . 2013; p. 4.

Mathôt S. Pupillometry: Psychology, Physiology, and Function. Journal of Cognition. 2018 Feb; 1(1):16. http://www.journalofcognition.org/articles/10.5334/joc.18/, doi: 10.5334/joc.18.

Mathôt S, Fabius J, Van Heusden E, Van der Stigchel S. Safe and sensible preprocessing and baseline correction of pupil-size data. Behavior Research Methods. 2018 Feb; 50(1):94–106. http://link.springer.com/10.3758/s13428-017-1007-2, doi: 10.3758/s13428-017-1007-2.

Mathôt S, Grainger J, Strijkers K. Pupillary Responses to Words That Convey a Sense of Brightness or Darkness. Psychological Science. 2017 Aug; 28(8):1116–1124. http://journals.sagepub.com/doi/10.1177/0956797617702699, doi: 10.1177/0956797617702699.

Miller SD, Thompson S. Edge-Light pupil cycle time. British Journal of Ophtalmology. 1978 Jul; 62:495–500. doi: 10.1136/bjo.62.7.495.

Pearson J. New directions in mental imagery research. Association for Psychological Science. 2014; 23(3):178–183. doi: 10.1177/0963721414532287.

Pearson J. The human imagination: the cognitive neuroscience of visual mental imagery. Nature Reviews Neuroscience. 2019 Oct; 20(10):624–634. https://www.nature.com/articles/s41583-019-0202-9, doi: 10.1038/s41583-019-0202-9.

Pearson J, Clifford CWG, Tong F. The Functional Impact of Mental Imagery on Conscious Perception. Current Biology. 2008 Jul; 18(13):982–986. https://linkinghub.elsevier.com/retrieve/pii/S0960982208007239, doi: 10.1016/j.cub.2008.05.048.

Pearson J, Naselaris T, Holmes EA, Kosslyn SM. Mental Imagery: Functional Mechanisms and Clinical Applications. Trends in Cognitive Sciences. 2015 Oct; 19(10):590–602. https://linkinghub.elsevier.com/retrieve/pii/S1364661315001801, doi: 10.1016/j.tics.2015.08.003.

Roberts R, Callow N, Hardy L, Markland D, Bringer J. Movement Imagery Ability: Development and Assessment of a Revised Version of the Vividness of Movement Imagery Questionnaire. Journal of Sport and Exercise Psychology. 2008 Apr; 30(2):200–221. https://journals.humankinetics.com/view/journals/jsep/30/2/article-p200.xml, doi: 10.1123/jsep.30.2.200.

Sheehan PW. A shortened form of Betts' questionnaire upon mental imagery. Journal of Clinical Psychology. 1967 Jul; 23(3):386–389. https://onlinelibrary.wiley.com/doi/10.1002/1097-4679(196707)23:3<386::AID-JCLP2270230328>3.0.CO;2-S, doi: 10.1002/1097-4679(196707)23:3<386::AID-JCLP2270230328>3.0.CO;2-S.

Theodossiadis PG, Achtsidis V, Theodoropoulou S, Tentolouris N, Komninos C, Fountas KN. The Effect of Alpha Antagonists on Pupil Dynamics: Implications for the Diagnosis of Intraoperative Floppy Iris Syndrome. American Journal of Ophthalmology. 2012 Apr; 153(4):620–626. https://linkinghub.elsevier.com/retrieve/pii/S0002939411007719, doi: 10.1016/j.ajo.2011.09.030.

Thomas NJT, Edward NZ, editor, Mental Imagery. Metaphysics Research Lab, Stanford University; 2021. https://plato.stanford.edu/archives/fall2021/entries/mental-imagery/.

Wilbrand DH. Ein Fall yon Seelenblindheit und IIemianopsie mit Sectionsbefund. Dtsch Z Nervenheilk. 1892; p. 27.

Willander J, Baraldi S. Development of a new Clarity of Auditory Imagery Scale. Behavior Research Methods. 2010 Aug; 42(3):785–790. http://link.springer.com/10.3758/BRM.42.3.785, doi: 10.3758/BRM.42.3.785.

Zeman A, Dewar M, Della Sala S. Lives without imagery – Congenital aphantasia. Cortex. 2015 Dec; 73:378–380. https://linkinghub.elsevier.com/retrieve/pii/S0010945215001781, doi: 10.1016/j.cortex.2015.05.019.

Zeman AZJ, Della Sala S, Torrens LA, Gountouna VE, McGonigle DJ, Logie RH. Loss of imagery phenomenology with intact visuo-spatial task performance: A case of 'blind imagination'. Neuropsychologia. 2010 Jan; 48(1):145–155. https://linkinghub.elsevier.com/retrieve/pii/S0028393209003418, doi: 10.1016/j.neuropsychologia.2009.08.024.

Appendix 1

Questionnaires

Questionnaire upon Mental Imagery

Participants should use the following notation system:

1: I think of it, but do not have an image before me.
2: Very vague and hardly recognizable.
3: Vague and unclear .
4: Not so clear and vivid but still recognizable .
5: Generally clear and vivid.
6: Vivid and almost as clear as in reality.
7: Very vivid and clear as in reality.

Visualize a mental image of a friend you see on a regular basis. How clear and vivid do you see		
1.	The characteristic lines of the face, head, shoulders and upper body?	
2.	The most characteristic positioning of the head and the body?	
3.	The characteristic behavior while walking / stepping (e.g., stride length)?	
4.	The various colors of clothes that are worn often?	
5.	How clear and vivid is the visual image you hold before your eyes when you think of a setting sun?	

How clear and vivid can you imagine the taste of:		
21.	Salt?	
22.	White crystal sugar?	
23.	Apple juice?	
24.	Jam?	
25.	Your favorite dish?	

How clear and vivid can you imagine the sound when you think of:		
6.	The whistle of a train?	
7.	The horn of a car?	
8.	The meowing of a cat?	
9.	The sound of escaping steam?	
10.	The clapping of hands of applause?	

How clear and vivid can you imagine the smell of:		
26.	A stuffy room?	
27.	Boiling vegetables?	
28.	Roasting meat?	
29.	Freshly applied paint?	
30.	Leather?	

How clear and vivid can you imagine the touch of:		
11.	Sand?	
12.	Linen?	
13.	Fur?	
14.	A pinprick?	
15.	A luke-warm bath?	

How clear and vivid can you imagine the feeling / emotion of:		
31.	Fatigue?	
32.	Hunger?	
33.	A sore throat?	
34.	Sleepiness?	
35.	Satisfaction after a tasty meal?	

Think of what you do with your arms, legs, lips, etc. How clear and vivid can you imagine what you do when you:		
16.	Climb the stairs?	
17.	Jump across a stream?	
18.	Draw a circle on a piece of paper?	
19.	Reach for an object on a high shelf?	
20.	Kick something away?	

Appendix 1 Figure 1. Questions presented in the QMI

Our questionnaire is accessible at the following link: https://enquetes-screen.msh-alpes.fr/index.php/927978?lang=en

Vividness of Visual Imagery Questionnaire

Participants should use the following notation system:

1: No image at all, you only "know" that you are thinking of an object.

2: Vague and dim.

3: Moderately clear and vivid.

4: Clear and reasonably vivid.

5: Perfectly clear and as vivid as normal vision.

In answering items 1 to 4, think of some relative or friend whom you frequently see (but who is not with you at present) and consider carefully the picture that comes before your mind's eye.		
1.	The exact contour of face, head, shoulders and body.	
2.	Characteristic poses of head, attitudes of body etc.	
3.	The precise carriage, length of step, etc. in walking.	
4.	The different colors worn in some familiar clothes.	

Think of the front of a shop which you often go to. Consider the picture that comes before your mind's eye.		
9.	The overall appearance of the shop from the opposite side of the road.	
10.	A window display including colors, shape and details of individual items for sale.	
11.	You are near the entrance. The color, shape and details of the door.	
12.	You enter the shop and go to the counter. The counter assistant serves you. Money changes hands.	

Visualize the rising sun. Consider carefully the picture that comes before your mind's eye.		
5.	The sun is rising above the horizon into a hazy sky.	
6.	The sky clears and surrounds the sun with blueness.	
7.	Clouds. A storm blows up, with flashes of lightening.	
8.	A rainbow appears.	

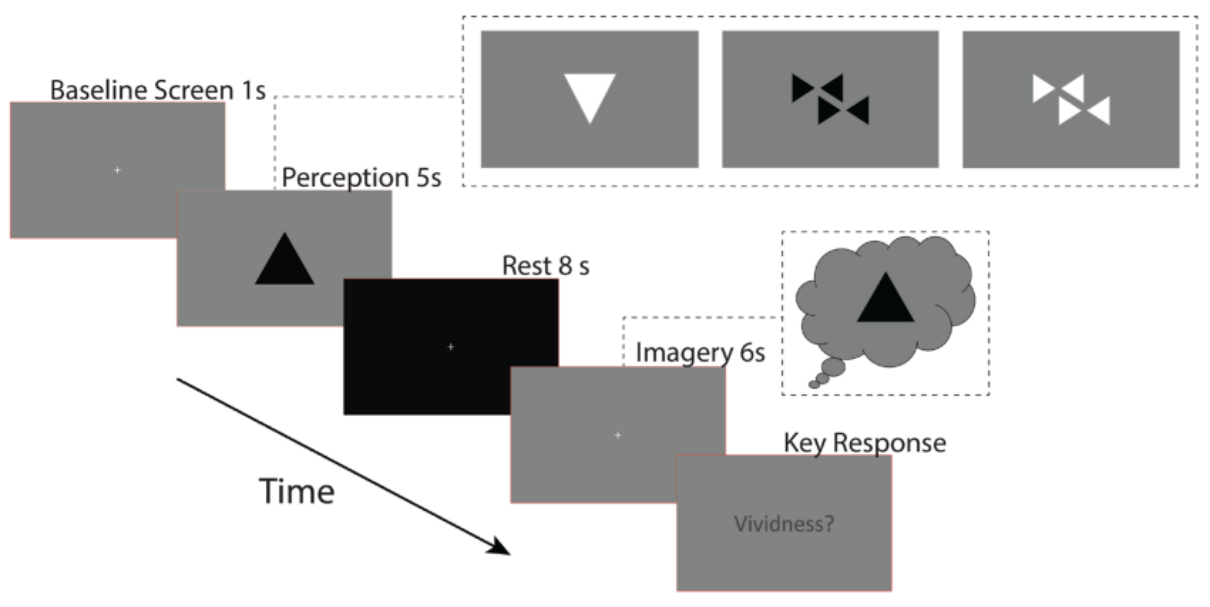
Finally, think of a country scene which involves trees, mountains and a lake. Consider the picture that comes before your mind's eye.		
13.	The contours of the landscape.	
14.	The color and shape of the trees.	
15.	The color and shape of the lake.	
16.	A strong wind blows on the tree and on the lake causing waves.	

Appendix 1 Figure 2. Questions presented in the VVIQ

Appendix 2

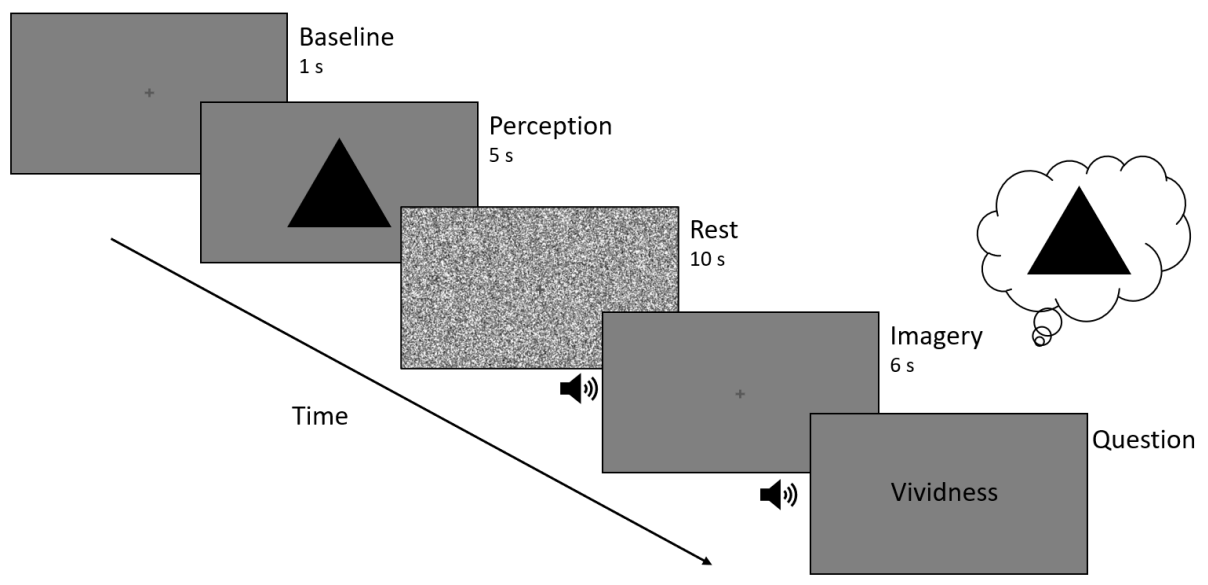
Protocols

Kay main experiment



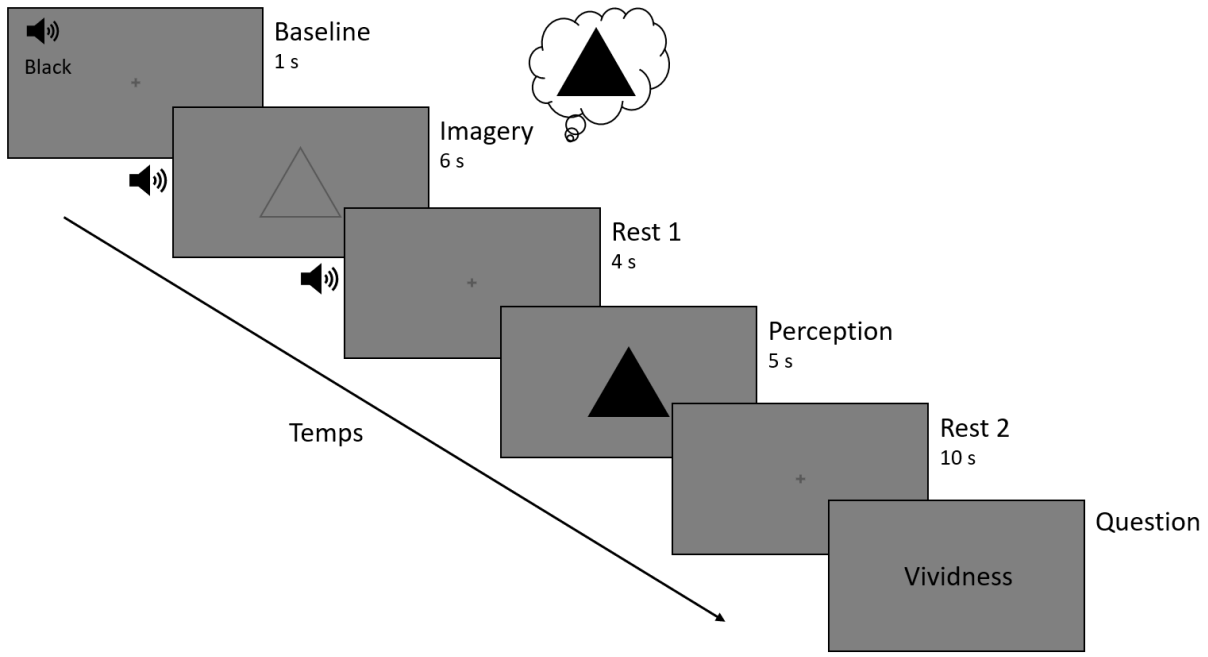
Appendix 2 Figure 1. Diagram describing the protocol of one trial of Kay and colleagues main experiment. In their experiments, they had one more condition. Stimuli were either "Set-Size- One" or "Set-Size-Four". (Extracted from Kay et al. (2022), Figure 1A

Perception first experiment



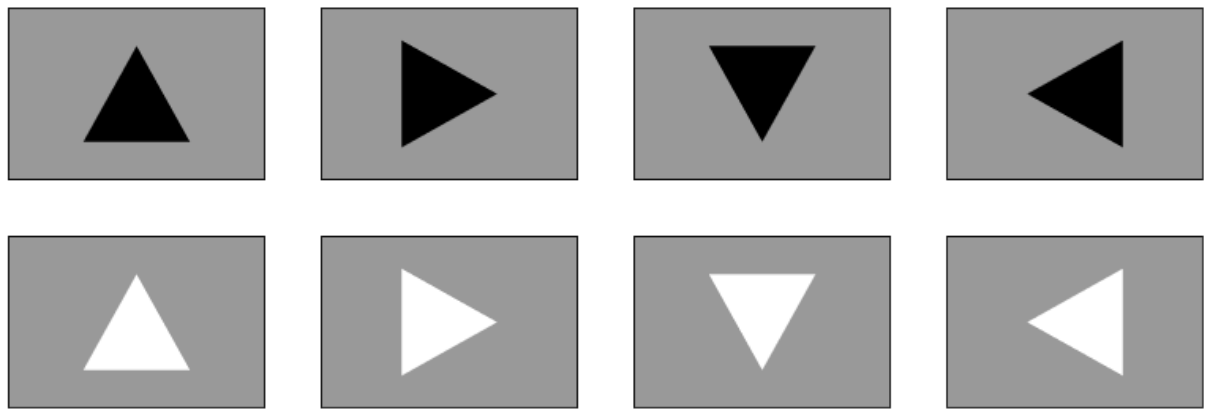
Appendix 2 Figure 2. Diagram describing the protocol of one trial of the perception first (PF) experiment. After participants answer to the vividness question, they will automatically start a new trial.

Imagery first experiment



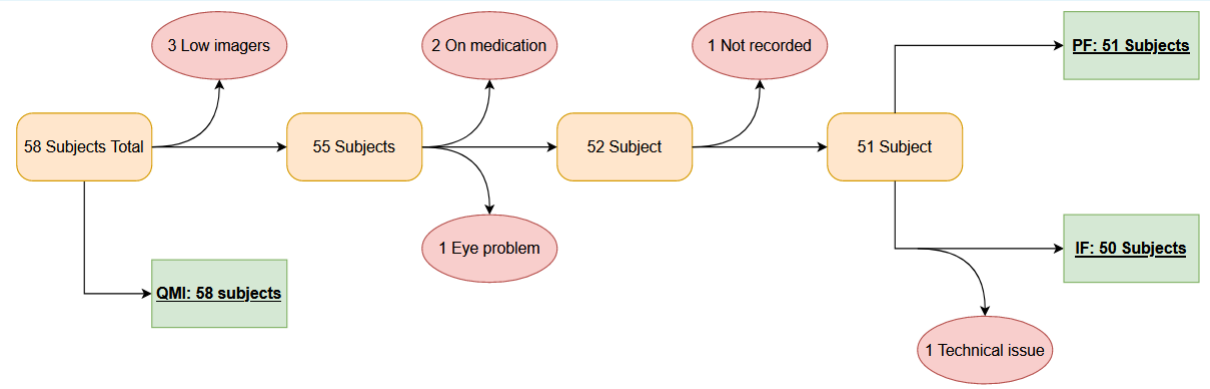
Appendix 2 Figure 3. Diagram describing the protocol of one trial of the imagery first (IF) experiment. After participants answer to the vividness question, they will automatically start a new trial.

Stimuli



Appendix 2 Figure 4. List of the triangles used during PF and IF experiments.

Exclusion tree



Appendix 2 Figure 5. Exclusion tree of our different experiments.

List of Figures

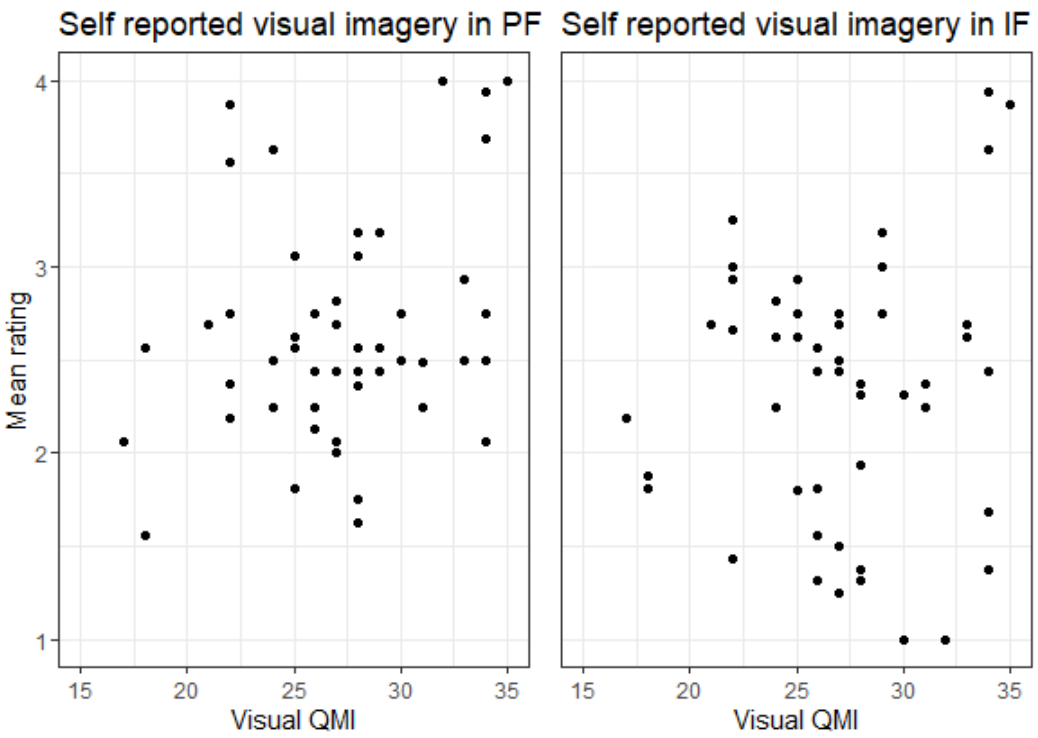


Figure 1–Figure supplement 1. Comparison of self reported visual imagery in the case of the PF experiment (N = 51 participants) and in the case of the IF experiment (N = 50 participants) against the visual QMI score. Each point represents one participant.

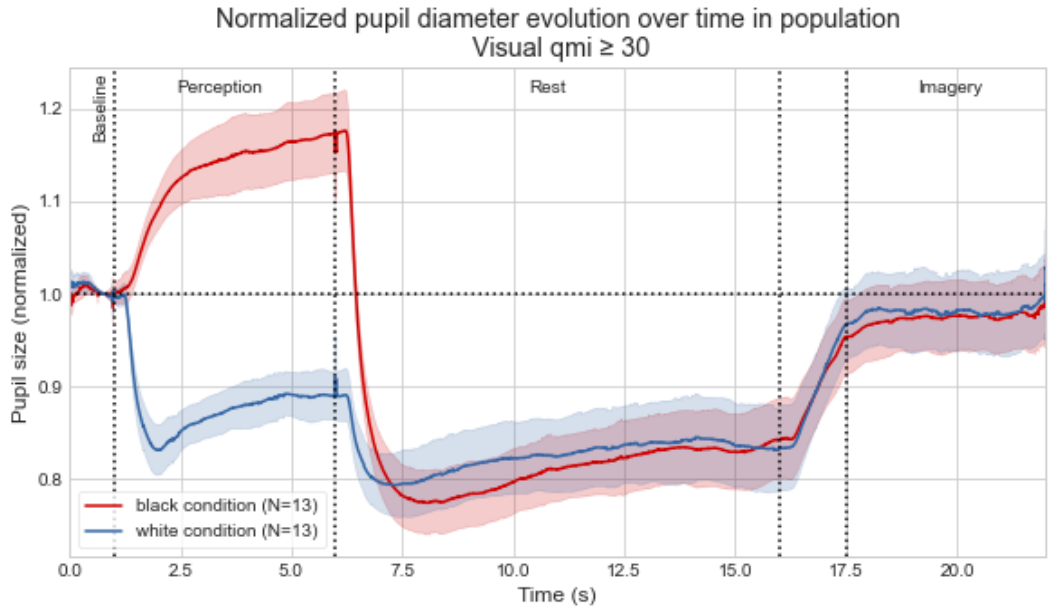


Figure 2–Figure supplement 1. Mean pupil size waveforms of our population sample with a visual QMI score greater than or equal to 30 (N = 13 participants). The lines represent the mean value and the shadows represent the standard error.

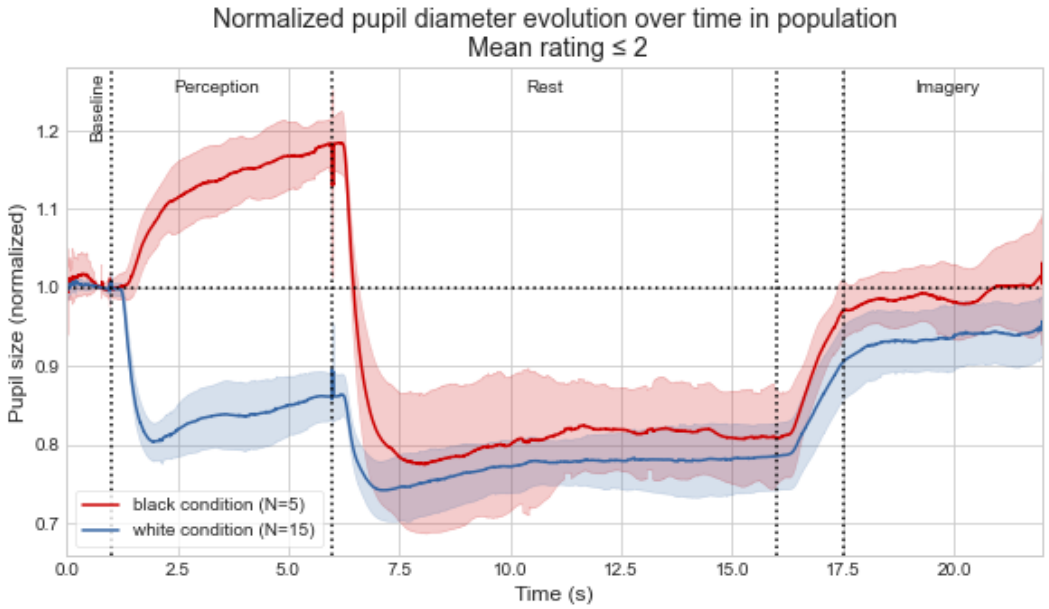


Figure 2–Figure supplement 2. Mean pupil size waveforms of our population sample with a mean rating score less than or equal to 2 (N = 5 participants in black condition and N = 15 participants in white condition). The lines represent the mean value and the shadows represent the standard error.

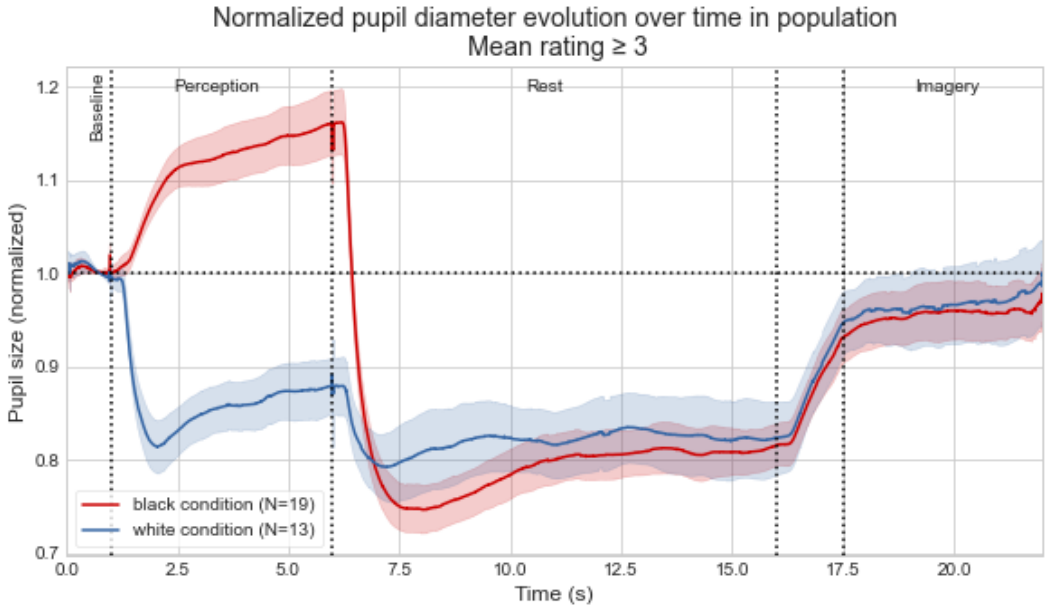


Figure 2–Figure supplement 3. Mean pupil size waveforms of our population sample with a mean rating of greater than 3 (N = 19 participants in black condition and N = 13 participants in white condition). The lines represent the mean value and the shadows represent the standard error.

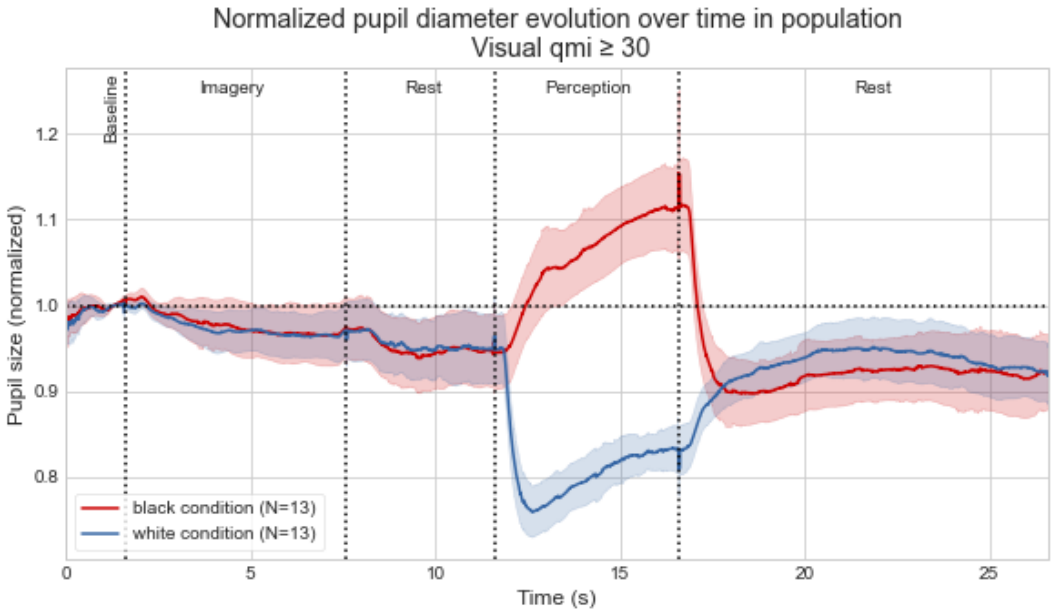


Figure 3–Figure supplement 1. Mean pupil size waveforms of our population sample with a visual QMI score greater than or equal to 30 (N = 13 participants). The lines represent the mean value and the shadows represent the standard error.

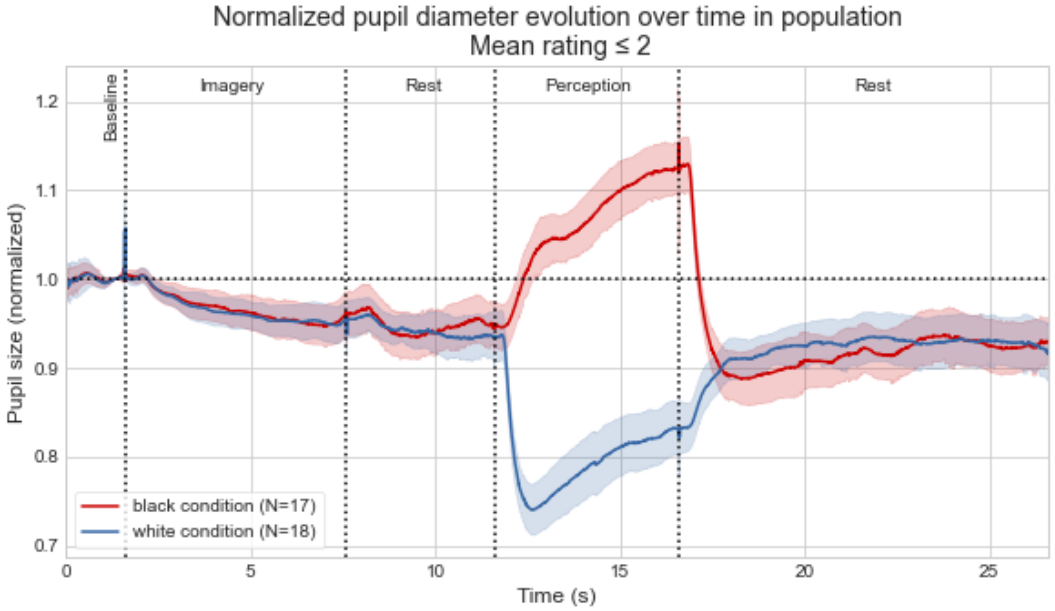


Figure 3–Figure supplement 2. Mean pupil size waveforms of our population sample with a mean rating score less than or equal to 2 (N = 17 participants in black condition and N = 18 participants in white condition). The lines represent the mean value and the shadows represent the standard error.

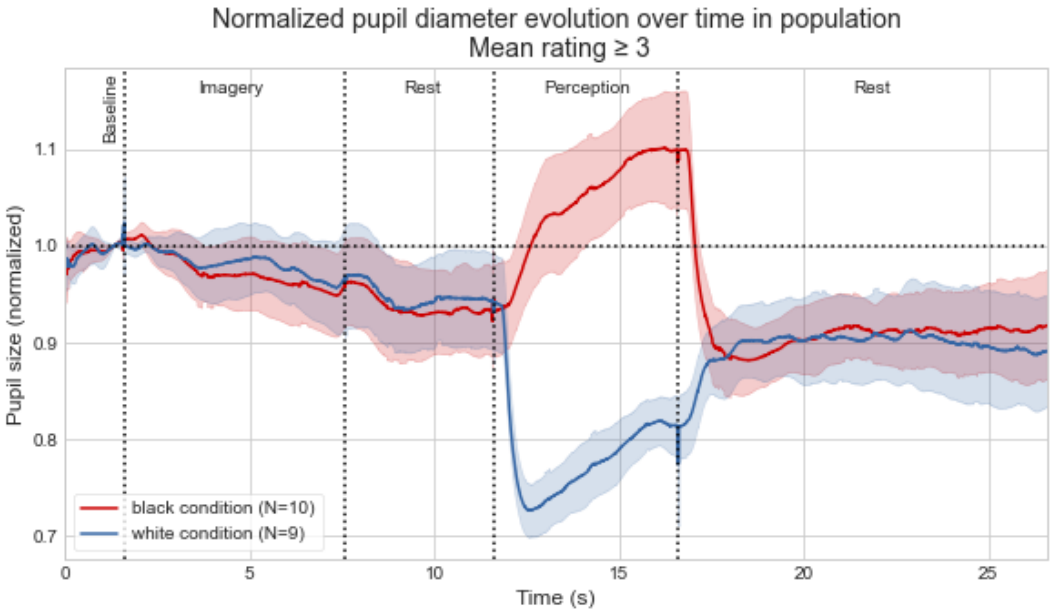


Figure 3–Figure supplement 3. Mean pupil size waveforms of our population sample with a mean rating of greater than 3 (N = 10 participants in black condition and N = 9 participants in white condition). The lines represent the mean value and the shadows represent the standard error.