

The eyes have it: The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength

Lachlan Kay^{1*}, Rebecca Keogh^{1,2*□}, Thomas Andrillion^{1,3} & Joel Pearson¹

¹ School of Psychology, University of New South Wales, Sydney, Australia

² Department of Cognitive Sciences, Macquarie University, Sydney, Australia

³ Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS

* These Authors contributed equally to this work

□ Corresponding Author

Abstract

The pupillary light response is an important automatic physiological response that optimises light reaching the retina. Recent work has shown that the pupil also adjusts in response to illusory brightness and a range of cognitive functions, however, it remains unclear what exactly drives these endogenous changes. Here we show that the imagery pupillary light response correlates with objective measures of sensory imagery strength. Further, the trial-by-trial phenomenological vividness of visual imagery is tracked by the imagery pupillary light response. We also demonstrated that there was no evidence for an imagery pupillary light response in a group of individuals without visual imagery (aphantasia), however, they did show perceptual pupil light responses and pupil dilation with larger cognitive load. Our results provide evidence that the pupillary light response indexes the sensory strength of visual imagery and also provides the first physiological validation of aphantasia.

Introduction

Our pupil's ability to change size is an important physiological response that adjusts the amount of light hitting the retina to optimise vision and protect the retina. Pupils constrict in response to brightness whereas they dilate in response to dark conditions (known as the pupillary light response or reflex); while these responses are related, they are considered to be driven by different neural pathways (see (Mathôt, 2018) for a review). These involuntary pupil responses were once thought to be driven only by afferent visual stimulation, or automatic activation from emotional responses (Bradley et al., 2008; Partala & Surakka, 2003), however, recent studies suggest that pupil size is sensitive to higher order perceptual and cognitive processes. For example, subjective interpretation of equiluminant stimuli, grayscale images of the sun elicit greater pupil constriction than those of the moon (Binda et al., 2013b). The target of covert visual attention can drive pupillary light responses (Binda et al., 2013a), as can visual working memory content (Hustá et al., 2019, Zokaei et al., 2019), but see (Blom et al., 2016). Further, evidence suggests that it might be mental imagery that is driving some of these cognitively induced pupil responses (Laeng & Sulutvedt, 2014) and recent work has shown that there are pupillary light responses even when reading or listening to words conveying some level of brightness (Mathôt et al., 2017). Hence it remains unknown if the variations in pupil response to equiluminant stimuli are due to high-level semantic content or low-level visual imagery.

Visual imagery is considered a useful and often essential tool in many aspects of cognition. It plays an important role in the retrieval of items from short and long term memory (Pearson, 2019), visual working memory (Keogh & Pearson, 2011, 2014a; Pearson & Keogh, 2019), acquisition of language (Just et al., 2004), and spatial navigation (Sack et al., 2005). It is also used for simulating both past and potential future events (Schacter et al., 2012; Schacter &

Madore, 2016), the latter often as a form of self-motivation for goal attainment (Szpunar et al., 2007). As essential to cognition as it might appear, large individual differences exist in visual imagery and its vividness. Some people report imagery as so vivid it feels almost like perception, while a small percentage of otherwise healthy people seemingly do not have the capacity for visual imagery at all - they report that when they think about how an object looks, there is no sensory-like experience of it whatsoever (Galton, 1880). This condition has been recently termed ‘aphantasia’ (Zeman et al., 2015); it can be congenital, persisting throughout one’s lifetime (Zeman et al., 2015) or acquired (Zeman et al., 2010), is associated with a range of differences in general cognition (Dawes et al., 2020), including dampened fear responses to imagined scary scenarios (Wicken, Keogh & Pearson, 2021). The existence of aphantasia has also been established using objective techniques that measure the low-level sensory elements of imagery (Keogh & Pearson, 2018).

The rationale of the current study is to accurately and objectively utilise individual differences in mental imagery (both in the general population and aphantasia) to provide strong evidence that it is the sensory strength and subjective vividness of imagery that drives the cognitive pupillary light response. Similar rationale has been previously used by linking the vividness and objective sensory strength of imagery to behavioural or neurological measures (Bergmann et al., 2016; Shine et al., 2015; Wassell et al., 2015). If imagery plays a causal role in endogenous pupil size changes, then individual differences in imagery should be reflected in these measures.

Here we utilised both subjective and objective measures of visual imagery ability and show that, within the same individual, greater pupillary light responses during imagery are associated with reports of stronger and more vivid imagery. We then used this task to compare

imagery strength between individuals and test the veracity of the self-reported lack of imagery in aphantasia. We show that while aphantasic individuals display pupil contraction to perceptual brightness and dilation with effort (cognitive load), they do not show any evidence of pupil change in response to attempts at imagery– providing the first objective physiological evidence confirming the existence of aphantasia.

Results

The imagery pupillary light response in the general population

In the pupillometry imagery task (based on (Laeng & Sulutvedt, 2014); see figure 1A), participants who reported having visual imagery were presented with one or four ‘Bright’ or ‘Dark’ triangles for 5s (see supplementary figure S1 for images used). Following this they viewed a blank screen for 8 seconds (which allowed any after-images to fade) and were then instructed to imagine the prior image/s for six seconds, after which they rated the vividness of their imagery from 1-4. Pupils showed a clear pupillary light response from perceptual images (Figure 1B; perception section; a RM mixed ANOVA found a significant effect of perceptual luminance $F(1,41) = 190.02, p < .001, \eta^2 = .77$) This trend was mirrored in the imagery period showing a significant main effect of imagery luminance (Figure 1B box insets: imagery section; RM mixed ANOVA, $F(1, 41) = 59.24, p < .001, \eta^2 = .30$), indicating that imagery also demonstrates a pupillary light response. Post-hoc analysis using the Bonferroni correction for multiple comparisons found that for both Set-Size-One and Set-Size-Four, the pupil size in the Dark condition was significantly greater than in the Light condition during imagery (both $p < .001$, see Figure 1C). There was no main effect of set-size during perception $F(1,41) = 2.67, p = .11, \eta^2 = .01$. However, there was a significant main effect of set-size during imagery $F(1, 41) =$

64.91, $p < .001$, $\eta^2 = .23$ with less constriction/more dilation for set-size four (when averaged across the brightness conditions). This is consistent with previous studies suggesting that pupil size is influenced by cognitive load (Kahneman & Beatty, 1966; Laeng et al., 2011; van der Wel & van Steenbergen, 2018).

Prior behavioural work suggests we have reasonable metacognition of visual imagery, i.e. we are able to estimate the strength of imagery on a trial by trial basis (Pearson et al., 2011; Rademaker & Pearson, 2012). Here we compared pupil responses to the trial-by-trial ratings of vividness. Pupil difference scores are shown as a function of intra-individual vividness ratings for set-size-one and set-size-four (see Figure 1D). A 2 x 4 linear mixed-effects analysis (2 (set size: one and four), x 4 (vividness rating: 1, 2, 3, 4)) demonstrated there was a significant effect of vividness ($\chi^2(3) = 49.54$, $p = 1.004e^{-10}$), with larger pupillary light response for more vivid imagery trials (for both set sizes, see figure 1D and fixed effects estimates in Table S1). These data provide novel evidence individuals can reliably evaluate the comparative vividness of single episodes of imagery. Further, these data demonstrate that the pupillary light response also tracks the phenomenological vividness of visual imagery from moment to moment.

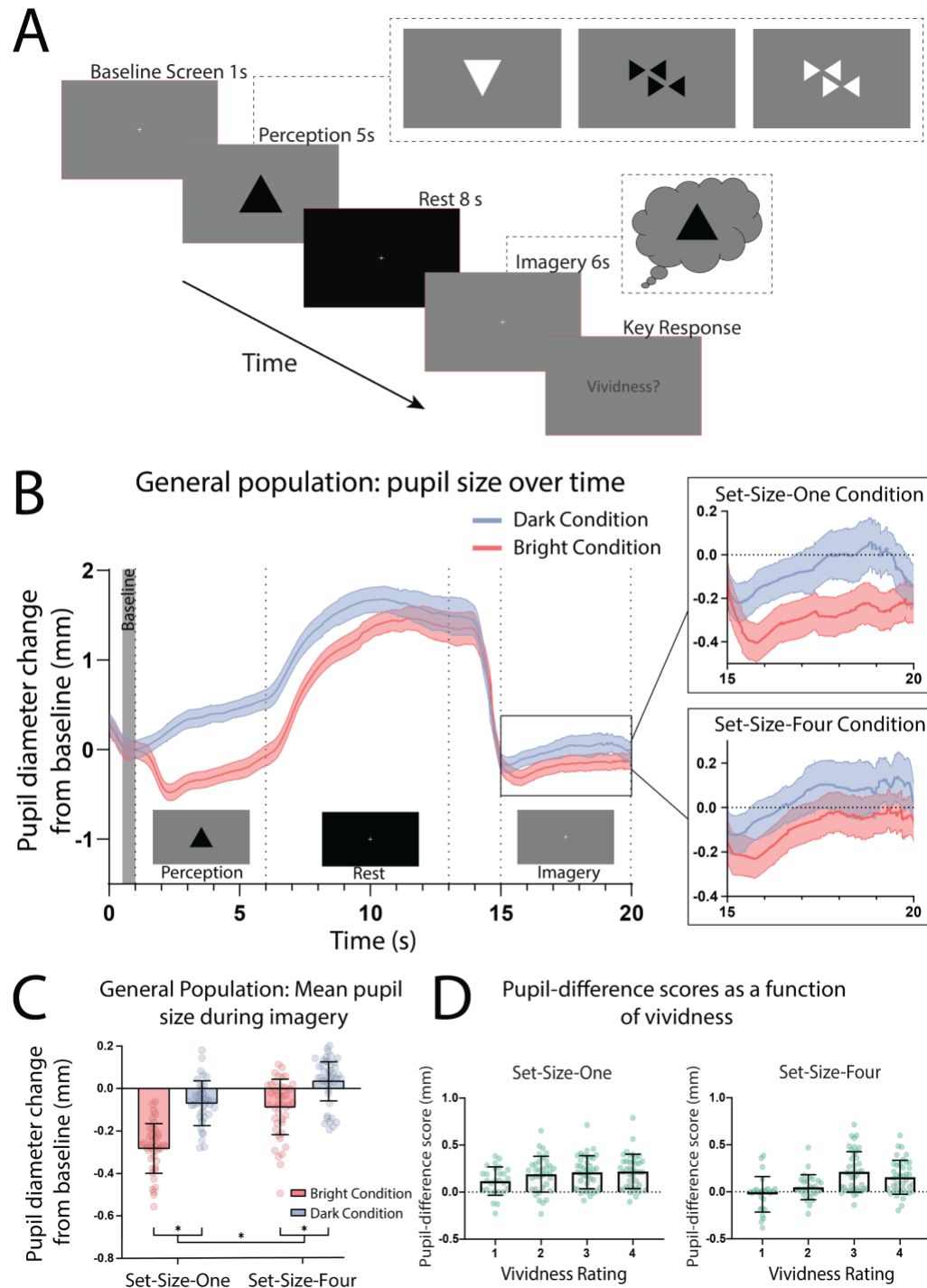


Figure 1. Pupillary response task schematic and eye-tracker results for the general population. (A) Pupillometry imagery experiment timeline. Each trial began with the presentation of a white fixation cross at the center of a grey screen (baseline) for 1 s. An image was then presented at the center of this grey screen for 5 s (either one or four triangles of varying brightness, see supplementary materials Figure S1 for illustrations of all stimuli). Participants were instructed to focus on the stimuli during this time and memorise its size, orientation and level of brightness. Next, a black screen with a white fixation cross was

presented for 8 s, allowing the perceived after-image to completely fade and pupils to dilate back to equivalent resting levels. The grey baseline screen was then presented again for 6 s. During this time, participants were cued (via two auditory beeps) to actively imagine the stimuli observed previously during that trial. Lastly, participants were prompted to report the vividness of their imagery during those previous 5 s on a scale of 1-4 (1 being “not vivid at all – no shape appeared in imagery”; 4 being “very vivid – almost like seeing it”) via key response. (B) Mean pupil size waveforms for the general population, presented as mm change from baseline. Left panel: data averaged across the course of a trial for Bright (red lines) and Dark (blue lines) conditions for the general population. *Right Panels:* set-size-one and set-size-four conditions are shown separately during the imagery period (i.e., pupil size from second 15 to second 20). Shaded error bands represent the standard error of the mean (\pm SEM). (C) Mean pupil size change from baseline during imagery (i.e., averaged from seconds 15 - 20 of trials) of Bright (red bars) and Dark stimuli (blue bars). (D) Pupil-difference scores (difference in pupil size during imagery between bright and dark conditions) as a function of subjective vividness ratings for set-size-one and set-size-four conditions. Data points represent one participant. Error bars indicate \pm SEM, calculated across participants. .

If the sensory strength of imagery is indeed driving the imagery pupillary light response, then the degree to which this response occurs should be related to independent objective measures of imagery strength in each individual. To assess this, we utilised the binocular rivalry method (Pearson, 2014; Pearson et al., 2008), which allows the objective assessment of the sensory strength of imagery, without relying on any subjective reports (Chang & Pearson, 2018). This is achieved by measuring the degree to which an individual’s imagery biases subsequent binocular rivalry perception. We compared pupil-difference scores (imagery of dark stimuli–bright stimuli, such that larger scores indicate a larger pupillary light response) with imagery strength measured using the binocular rivalry paradigm, in which higher priming scores indicate stronger imagery (Figure 2A; (Pearson et al., 2008, 2011)). Within the general population, degree of pupil change in the set-size-one condition correlated positively with imagery strength, using Pearson’s correlation coefficient ($r_p(41) = .62, p < .001$, see Figure 2B: green circles and green trendline). The set-size-four pupil data set violated normality (Shapiro-Wilk test, $p = .003$), therefore, the Spearman’s correlational coefficient was used to assess its relationship with binocular rivalry priming. A significant positive correlation was found between set-size-four

pupil-difference scores and binocular rivalry priming ($r_s(41) = .46, p = .002$, see figure 2C: green circles and green trendline). This provides further evidence that the sensory strength of imagery content is driving the imagery pupillary light response.

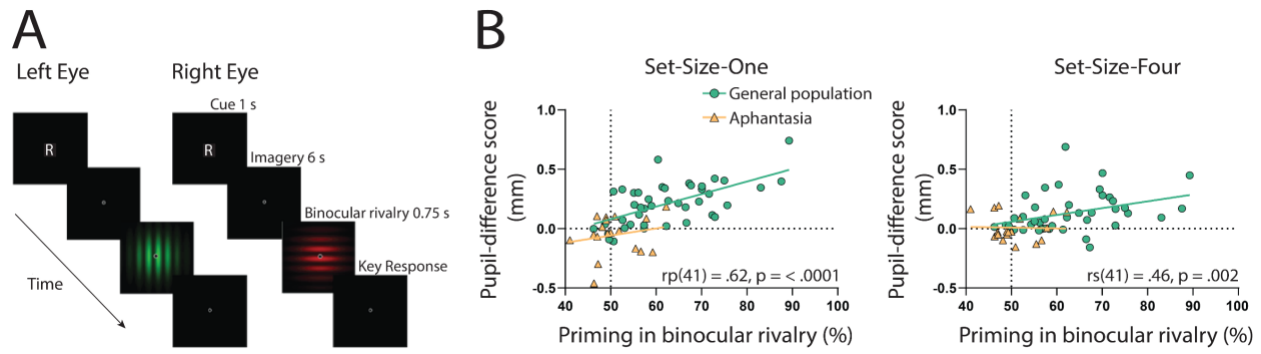


Figure 2. *Binocular rivalry task schematic and correlational results.* (A) Example of an imagery trial for the binocular rivalry paradigm. Participants were cued to imagine either a red or green Gabor pattern prior to binocular rivalry with the letter ‘R’ or ‘G’ (750ms). Participants then imagined the image for 6s, after which they were presented with the binocular rivalry display (750ms) and were asked to indicate which image was dominant. Trials where participants reported seeing the pattern they were cued to imagine as dominant were denoted as ‘primed’ trials. The number of primed trials divided by the total number of trials (excluding mock trials and mixed percepts) was used to calculate a percent primed score for each participant. (B) Correlation between visual imagery strength, as measured by the pupillary response task (pupil-difference score: difference between bright and dark conditions) and visual imagery strength as measured by the binocular rivalry task. Set-size-one (left) and set-size-four (right) conditions are shown. Scatterplots show the general population (green circles and green trendline) and aphantasic individuals (yellow triangles and yellow trendline) data. Correlation coefficients refer to the general population only (green trendline). All data points represent one participant.

Aphantasia and the imagery pupillary light response

Our results indicate that the strength of the content of imagery drives the imagery pupillary light response in participants who experience visual imagery. The involuntary nature of this response provides a valuable objective measure of imagery strength. Accordingly, we sought to utilise this finding to test the veracity of a condition called aphantasia, i.e., if they truly lack visual imagery, they should not show a pupillary light response to imagined images. However, if

aphantastic individuals do show an imagery based pupillary light response, one might interpret this as a form of imagery existing, but below threshold for conscious phenomenological awareness. We ran this same study in 18 aphantastic participants and compared their performance to that of the general population. These participants had contacted the lab reporting their lack of visual imagery and asked to participate in our research. They were also unaware of the goals and hypotheses of the current study. Aphantasia was confirmed in these individuals using self-report questionnaires (Vividness of Visual Imagery Questionnaire (VVIQ) score < 32) and by means of our binocular rivalry priming method (priming < 65%), based on cut-off points used in previous research (Keogh & Pearson, 2018).

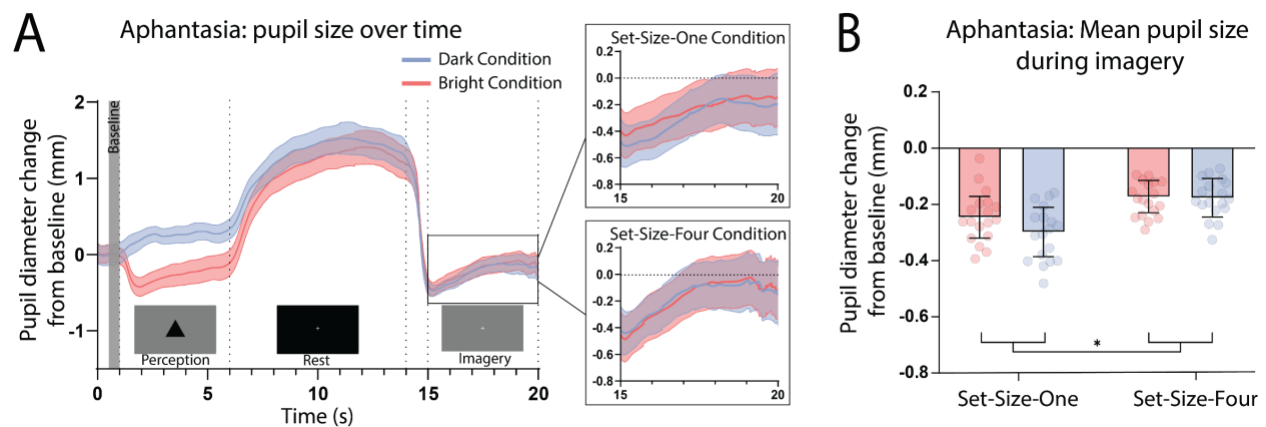


Figure 3. Pupillary response eye-tracker results for the aphantastic population. (A) Mean pupil size waveforms over time. *Left panel:* data averaged across the course of a trial for Bright (red lines) and Dark (blue lines) conditions for the general population. *Right Panels:* set-size-one and set-size-four conditions are shown separately during the imagery period. (B) Mean pupil size change from baseline during imagery (i.e., averaged from seconds 15 - 20 of trials) of Bright (red bars) and Dark stimuli (blue bars). Error bars indicate \pm SEM, calculated across participants. * $P < 0.05$.

Here we again found a strong effect of luminance in the perceptual phase of the task for the aphantastic participants (Figure 3A: perception section; $F(1,17) = 81.18$, $p < .001$, $\eta^2 = .70$), reflecting a functional pupillary light response. However, we found no significant effect of

luminance on pupil size during imagery (Figure 3A box insets: imagery section; $F(1, 17) = 1.02$, $p = .33$, $\eta^2 = .02$ and Figure 3B shows the lack of pupil diameter change for bright stimuli (red bars) and dark stimuli (blue bars)). Similarly, to the general population, there was no main effect of set-size during perception $F(1,17) = 1.92$, $p = .18$, $\eta^2 = .01$ however there was a significant main effect of set-size during imagery $F(1, 17) = 36.01$, $p < .001$, $\eta^2 = .27$, with greater pupil diameters for set-size four compared to set-size one (when averaged across the brightness conditions). This suggests that the aphantasic participants were actively engaging in the imagery task and exerting greater cognitive effort for the larger set-size (van der Wel & van Steenbergen, 2018). In comparison to the general population, 61.11% (11/18) of the aphantasic individuals had difference scores that were lower than or equal to 0 for set size one as compared to 9.5% (4/42) of the general population (see figure 2B). To confirm this absence of an imagery effect in the aphantasia population, we compared the pupil-difference score obtained when comparing the bright and dark condition for the control and aphantasia groups, and computed a Bayes Factor (H_0 : score=0; H_1 : score \neq 0; see Methods). Controls showed very strong evidence for H_1 ($BF_{10} > 10^{10}$; Bayesian one-sample t-test) whereas the aphantasia population showed evidence for the null effect ($BF_{01} = 3.180$). A direct comparison between the control and aphantasia group using a Bayesian repeated measure ANOVA (see Methods) showed very strong evidence for an effect of group ($BF_{10} > 10^6$). Finally, and as expected, pupil difference scores (imagery of dark stimuli–bright stimuli) did not significantly predict imagery strength (measured using the binocular rivalry paradigm) for the aphantasic population (Figure 2B: blue triangles; set-size-one: $r_p(17) = 0.20$, $p = 0.44$; set-size-four: ($r_p(17) = -0.08$, $p = 0.76$). It should be noted that we could not perform an analysis on the vividness data in the same way as was done with the general population (figure 1D) as the aphantasic individuals did not have any variation in their

vividness ratings, reflecting their lack of subjective visual imagery (see supplementary figure S4).

Age disparities between the groups are a potential confounding variable. This factor is of particular importance because the sensitivity of the pupillary light response, as well as maximum pupillary constriction velocity and acceleration, are thought to decline with age, beginning at 40 - 50 years old (Fotiou et al., 2007; Lobato-Rincón et al., 2014). However, trial time-course pupil waveforms are very similar for both general and aphantasic populations (figure 1B and figure 3A respectively). Both groups exhibited similar levels of pupil change during the perception phase of the task. Furthermore, a two-way ANCOVA was run on pupil-difference scores between general population and aphantasic groups with age as a covariate. Levene's test and normality checks were carried out and the assumptions were met. We found a significant difference in pupil-difference scores ($F(1,117) = 5.23, p = .02, \eta^2 = .04$) between the groups when accounting for age. This provides evidence that decreased pupil responsiveness with age was not driving the lack of an imagery pupillary light response in the aphantasic population.

Another possible explanation of our findings could be that perception of the images, and lingering visual persistence and sluggish pupil responses could be driving our results. If this is the case, we would expect that pupil diameter during the perception of the images should correlate with pupil size during imagery for the corresponding images. Further, the pupillary light reflex during perception should be more pronounced in the control than the aphantasic populations. To investigate this possible alternative explanation of our data we first assessed the correlations between pupil diameter during perception of bright and dark images for set size one and four and their corresponding imagery conditions and found there were no significant correlations between any of the perception and imagery conditions, or the difference scores for

set size one and four (see supplementary figure S5). Next, we assessed whether the aphantasic individuals demonstrated any significant difference in their pupil responses to perceptual stimuli by running a 2 (image: bright & dark) x 2 (set size: 1 & 4) x 2 (group: aphantasic & controls) repeated measures ANOVA on the pupil diameter during the five-second perceptual period of the task (see figure 1A for task timeline). There was no main effect of imagery group $F(1,58) = 1.15, p = .29, \eta^2 = .01$ and no significant interactions between imagery groups and any other factor (see supplementary figure S6). These findings suggest the observed pupil responses during the imagery period of the task is unlikely to be a carry-over effect of the previous sensory response to perceived images.

Taken together these data from the general population and aphantasic individuals suggest that it is the content and ability to form vivid visual images, not the voluntary attempt to do so or the semantic content, that is driving the imaginary pupillary light response, providing the first evidence that these pupil changes are due to the sensory strength of imagery content and are not driven by higher-level semantic content.

Discussion

Our results provide novel evidence that our pupils respond to the vividness and strength of a visual image being held in mind, the stronger and more vivid that image the greater the pupillary light response. Our data provides the first evidence linking the pupil response to strength and vividness of imagery, not only between individuals, but also within an individual as imagery vividness fluctuates from moment-to-moment (Dijkstra et al., 2017; Pearson et al., 2011; Rademaker & Pearson, 2012). Finally, we show that, as a group, there is no evidence of this pupil response in individuals without mental imagery (aphantasia).

How might the content of mental imagery be driving the pupillary light response? One interpretation of these findings is that this imagery pupillary response is a by-product of the top-down modulation of midbrain-level visual circuitry (pretectal olivary nucleus, superior colliculus; (Joshi & Gold, 2020)), which occurs when imaging vividly, resulting in these regions interpreting this modulation as coming from external or afferent stimuli, and responding accordingly (Larsen & Waters, 2018; Schwalm & Jubal, 2017). In this case, the pupil would be responding to imagined luminance in much the same way that it responds to retina-based light sources. This is consistent with current data and models proposing shared mechanisms between visual imagery and perception (Dijkstra et al., 2017, 2019; Ganis et al., 2004; Naselaris et al., 2015; Xie et al., 2020) and the idea that visual imagery functions much like a weak version of afferent perception (Pearson, 2019), supporting the idea that the stronger or more vivid an individual's imagery is, the more 'perception like' their imagery is.

An alternative mechanistic account might be that pupil diameter is encoded along with the original visual information e.g., bright object, and hence is replayed during memory decoding to form the mental image. This would be in a similar manner to theories proposing a functional role of eye movements during imagery generation from memory (Wang et al., 2020). It will be up to future work to uncover the exact mechanistic account of imagery induced pupil changes.

Here we also provide the first objective physiological evidence of an extreme lack of visual imagery in aphantasic individuals. Aphantasia has largely been defined using subjective means (Dawes et al., 2020; Jacobs et al., 2018; Pounder et al., 2018; Zeman et al., 2015), but see (Keogh & Pearson, 2018). Accordingly, people have remained sceptical about its true nature and possible psychogenic basis (de Vito & Bartolomeo, 2016). Our data demonstrate that using a non-visual strategy (no imagery in aphantasia) to think about bright and dark objects does not

induce a pupillary light response. These data simultaneously provide strong evidence linking the pupillary light response to mental imagery, as well as supporting the behavioural work showing that aphantasic individuals indeed lack visual sensory imagery (Keogh & Pearson, 2018).

Because the pupillary light response is involuntary (Bouffard, 2019), we can consider these findings as an unbiased neurophysiological measure of aphantasia, further ruling out claims of a psychogenic basis (de Vito & Bartolomeo, 2016). Not only do these data show that pupillary light response can be an objective index of imagery strength in studies of imagery in general populations, it also provides a new low-cost objective measure for aphantasia that is uniquely based on a physiological mechanism and not reliant on self-report.

Could a lack of active engagement during imagery explain the aphantasia results? Put another way, are they refusing to imagine (de Vito & Bartolomeo, 2016)? We think this is highly unlikely as pupil size *did* increase as a function of set-size for aphantasic individuals, as has previously been shown in the general population, demonstrating the typical relationship between cognitive effort or arousal and pupil dilation (Kahneman & Beatty, 1966; van der Wel & van Steenbergen, 2018). This demonstrates active task engagement and that the aphantasic population tested here has functional pupil circuitry – ruling out abnormal pupillary responses as a confound or that they were simply ‘refusing’ to actively participate in the task due to demand characteristics or a belief that they are unable to imagine (de Vito & Bartolomeo, 2016). It also reveals that regardless of what imagery strategy aphantasic participants are implementing (e.g., propositional, spatial, language-like) to recall information about the shapes, they require greater cognitive effort to simultaneously maintain a larger number of shapes in their mind.

Another possible explanation of our results is that perceptual pupillary light responses are lingering throughout each trial and driving the observed imagery pupil response. If this is the

case, then pupil responses during perceptual viewing and imagery should be correlated, however we did not find any such correlations (see figure S5). Further, when directly comparing the perceptual pupil responses between the general population and aphantasic individuals, there was no main effect of group or interaction between group and stimuli brightness or set size (see supplementary figure S6). This demonstrates that there is no significant difference in the perceptual pupillary responses between the two groups, making it unlikely that aphantasic individual's lack of an imagery pupillary response is due to a lack of perceptual response. Finally, we also asked participants if they perceived any after images during the imagery period and any participants who reported they did were excluded from the study. Taken together, these results suggest that it is unlikely that the pupillary response to perceptually viewing the images is driving our observed imagery pupillary responses, and the lack thereof in the aphantasic individuals. Instead, it appears the pupillary light response during the visual imagery period reflects the wilful generation of imagery in the mind's eye of those who experience visual imagery. This is further substantiated by the strength of visual imagery (measured using the binocular rivalry paradigm) correlating with the imagery pupillary light reflex, but not the perceptual pupillary light reflex (see figure 2B and S7).

We also found that in the pupillary response imagery task, higher within-trial reports of vividness are reflected by greater pupillary light responses (within-subjects effects; see figure 1D). This indicates that participants were able to accurately evaluate the vividness of individual episodes of imagery in comparison to other vividness episodes on previous trials. However, average vividness ratings did not correlate with their pupil-difference scores, that is, participants who gave higher vividness ratings on average did not necessarily have increased pupil light responses in response to imagery (between-subjects effects; see supplementary materials Figure

S2). Participant's scores on the Vividness of Visual Imagery Questionnaire also did not correlate with pupil difference scores (between-subjects effects; see supplementary materials Figure S3). This suggests that participants might have difficulties in accurately reporting their strength of sensory visual imagery on an absolute scale (i.e., from 'no image' to 'as vivid as perception'), and brings into question the reliability of these subjective measures of imagery and highlights the utility of using objective or online (i.e. in a task), and less trait-like measures when studying visual imagery.

Recent studies have shown pupil size is also modulated by the content of visual working memory (Blom et al., 2016; Hustá et al., 2019; Zokaei et al., 2019). It is interesting to note here that previous work has shown that imagery has been implicated as one mnemonic that can be used to retain information in mind during visual working memory tasks (Albers et al., 2013; Keogh & Pearson, 2014a, 2017). This highlights the possibility that it is imagery, being used as a mnemonic strategy, that is driving the pupillary light response observed in visual working memory experiments (Pearson & Keogh, 2019). Although many participants report using a visual imagery strategy during these tasks, some participants report using a non-visual imagery strategy when remembering visual information, and recent work demonstrates that aphantasic individuals can perform traditional visual working memory tasks just as well as control populations (Keogh, Wicken, et al., 2021). Measuring the pupillary light response in aphantasic individuals, and those who report not using an imagery strategy, while performing classic visual working memory tasks may help to further elucidate these differences in cognitive strategy use in a more objective manner.

One limitation that is important to note here is that our aphantasic sample contained a relatively small sample (18 participants) due to the relative rarity of this condition. Further our

two samples were not age matched, which may have affected our results, however seeing as there was no difference between the two groups for the perceptual pupillary light response, we think this is unlikely to be driving our findings. Future studies should aim to replicate and extend these findings with a larger group of aphantasic individuals and age matched controls.

To conclude, the present study demonstrates that the pupillary light response can be used as a physiological index of individual differences in the sensory and phenomenological strength of visual imagery, including the lack of visual imagery – aphantasia. Combining this measure with the binocular rivalry paradigm in favour of subjective alternatives will increase the reliability and objectivity of imagery test batteries and may lead to the development of more congenial theories of the mind’s eye.

Materials and Methods

Participants

Fifty-six psychology students with a mean age of 19.8 years (range 18-31, 27 female) were recruited for the study and participated for course credit. We aimed to obtain analysable data from a minimum of 40 participants, which should be a large enough sample to identify a strong positive correlation between pupil dilation and imagery, which is what we would expect if imagery content were driving the previously observed imagery pupillary light response (g*Power effect size = .5, $\alpha = .05$, $\beta = .95$). Fourteen of these participants were excluded from data analysis for not meeting a priori criteria (see Exclusion Criteria below), leaving 42 participants in the final general population sample.

The aphantasic individuals come from a rare population and for this reason we did not run a specific power analysis but aimed to collect a minimum of 15 participants. We had nineteen aphantasic individuals agree to participate in the study with a mean age of 35.8 years (aged 18-54, 12 female). One of these individuals was excluded from data analysis for not meeting a priori criteria (see Exclusion Criteria below), leaving 18 in the final sample. These participants had all contacted the lab regarding their aphantasia and asked to participate in our research. They were all reimbursed \$20 AUD per hour for their participation. All participants had normal or corrected to normal vision (i.e., glasses or contacts). Both experiments were approved by the UNSW Human Research Ethics Advisory Panel (HREAP-C 3182).

Apparatus

Apparatus Stimuli in all experiments were presented on an LCD display monitor (Dell UltraSharp U2419H) with 60 Hz refresh rate and a 1920 x 1080 resolution. Luminance values of all stimuli were measured using a Konica Minolta chroma meter (CS-100A). Participants placed their chin on a chin rest throughout the experiment to maintain fixation at a distance of 57cm from the monitor 13 and to limit head movements. The tasks were performed in a blackened room to eliminate any possible fluctuations in ambient light.

In the **pupillary response task**, pupil sizes and eye movements were recorded using head mounted eye-tracking glasses (Pupil, Pupil Labs GmbH, Berlin, Germany) (Kassner et al., 2014). Pupil diameter of participants' right eye was continuously sampled at 200Hz throughout the task. A pupil detection 3D algorithm locates the dark pupil in the infrared illuminated eye camera image, thus recording capabilities are not compromised by an absence of room lighting. Pupil diameter is then scaled to millimetres (mm) based on mean anthropomorphic eyeball diameter and corrected for perspective. The algorithm does not depend on corneal reflection, and is compatible with users who wear contact lenses and most eyeglasses (Kassner et al., 2014).

A second camera mounted on the glasses continuously recorded participants' field of view. Footage from this camera was subsequently assessed to ensure fixation on the computer monitor was maintained throughout the task. The experiment was designed using MATLAB (version R2017b). ZeroMQ plug-ins were used for cross-communication between eye-tracking and stimulus presentation platforms (Akgul, 2013). Pupil data was recorded with Pupil Capture v.1.10.20 (Pupil Labs) installed on an ASUS (GL502V) PC (Windows 10).

In the **binocular rivalry task**, participants wore red-green anaglyph glasses to ensure rivalrous stimuli were presented to left and right eyes in isolation. Responses of 1, 2 or 3 on a keyboard were used by participants to indicate which image dominated their perception during binocular rivalry (1 for green; 3 for red; 2 for perceptually mixed green and red).

Stimuli

For the **pupillary response task**, 32 achromatic shape stimuli were created for participants to perceive and then later imagine in their absence, across 32 trials. The stimuli were evenly divided based on a 2x2 factorial design, belonging to one of two luminance conditions ('Bright' or 'Dark') and one of two set-size conditions ('set-size-one' or 'set-size-four'). Shapes belonging to the Bright condition were either white with a luminance of 117 cd/m² or light grey with a luminance of 65 cd/m². Shapes in the Dark condition were black (1 cd/m²) or dark grey (9cd/m²). Set-size one stimuli consisted of a single equilateral triangle with 12.5 cm sides, subtending 12.5 degrees of visual angle. Set-size-four stimuli consisted of an arrangement of 4 four smaller equilateral triangles with a total surface area and luminance equal to that of the corresponding set-size-one triangles (see FigureS1 for illustration of all stimuli). Stimuli were also uniquely orientated at either 0°, 90°, 180° or 270° (e.g., 4 set-size-one black triangles, each with a different orientation. See supplementary materials for examples all possible shape orientations). This ensured that all 32 stimuli were unique and participants and were encoding information about a new stimulus on each trial, therefore avoiding the use of long-term memory. Set-size-four stimuli therefore subtended either 10.8° or 18.9° of visual angle depending on their orientation (see supplementary materials Figure S1 for illustrations of all stimuli).

All stimuli were presented on a grey background screen with a luminance of 26 cd/m². This same level of background luminance was used during measurement of baseline and imagery phases. A fixation cross on a black background with a luminance of 1 cd/m² was presented during the resting phase of each trial. All stimuli were created in MATLAB, using the Psychophysics Toolbox 3 extensions (Brainard, 1997).

For the **binocular rivalry task**, sinusoidal luminance modulated Gabor patterns were used as rivalrous stimuli; vertical-green (CIE chromaticity coordinates: $x = .275$, $y = .590$) and horizontal-red (CIE chromaticity coordinates: $x = .492$, $y = .372$), both with a mean luminance of 8.35 cd/m² and 7.1 degrees of visual angle. In each trial, both patterns were presented at the same time around a fixation point at the centre of a black background screen. Mock rivalry stimuli (a single 15 Gabor pattern spatially divided into half vertical-green and half horizontal-red) were used on 12.5% of trials to measure the influence of decisional bias or lack of attention to the task. More details on the binocular rivalry task can be found in (Keogh & Pearson, 2014b).

Procedure

Pupillometry imagery experiment timeline: Each trial began with the presentation of a white fixation cross at the centre of a grey screen (baseline) for 1 s. An image was then presented at the centre of this grey screen for 5 s (either one or four triangles of varying brightness, see supplementary materials Figure S1 for illustrations of all stimuli). Participants were instructed to focus on the stimuli during this time and memorise its size, orientation and level of brightness. Next, a black screen with a white fixation cross was presented for 8 s, allowing the perceived after-image to completely fade and pupils to dilate back to equivalent resting levels. The grey

baseline screen was then presented again for 6 s. During this time, participants were cued (via two auditory beeps) to actively imagine the stimuli observed previously during that trial. Lastly, participants were prompted to report the vividness of their imagery during those previous 5 s on a scale of 1-4 (1 being “not vivid at all – no shape appeared in imagery”; 4 being “very vivid – almost like seeing it”) via key response.

Binocular rivalry paradigm: Participants were cued to imagine either a red or green Gabor pattern prior to binocular rivalry with the letter ‘R’ or ‘G’ (750ms). Participants then imagined the image for 6s, after which they were presented with the binocular rivalry display (750ms) and were asked to indicate which image was dominant (see Figure 2). Trials where participants reported seeing the pattern they were cued to imagine as dominant were denoted as ‘primed’ trials. The number of primed trials divided by the total number of trials (excluding mock trials and mixed percepts) was used to calculate a percent primed score for each participant.

Exclusion Criteria

Of the 56 participants recruited for the general population sample, 14 in total were excluded from data analysis due to not meeting a priori criteria.

Pupillary response task exclusions: 8 participants were excluded because more than 50% of their pupil data points were below the pupil detection algorithm confidence value of 0.6, provided by the Pupil Capture system. This cut-off point was derived prior to data collection and is the recommended cut-off point for obtaining accurate pupil size data (Pupil Labs). 3 participants were excluded due to reporting (during systematic post-task questioning) seeing after-images of the shape stimuli for longer than the 8 s black screen presentation (i.e., seeing after-images during the imagery phase of trials), because pupil size is known to be influenced by the induced compensatory light perception of an after-image (Tsujimura et al., 2003).

Binocular rivalry task exclusions: 3 participants were excluded due to having mock rivalry priming >66.67% (more than one incorrect response on the mock trials), which indicated either an influence of decisional bias or lack of attention to the task. An a priori cut-off point of scoring *both* below 65% priming on the binocular rivalry task and below 32 on the Vividness of Visual Imagery Questionnaire was used to exclude participants who potentially did not have visual imagery (i.e., may be aphantasic). No participants fell below this combined cut-off point thus none were excluded on this basis.

Of the 19 participants recruited for the aphantasic population, 1 was excluded from data analysis because more than 50% of their pupil data points were below the pupil detection algorithm confidence value of 0.6, given by the Pupil Capture system. All participants scored below both of the a priori cut-off points of 32 on the Vividness of Visual Imagery Questionnaire and 65% on the binocular rivalry task, therefore, no participants were excluded due to this criterion.

Data Analysis

For the pupillary response task, cubic spline interpolation was used to estimate pupil diameter during periods where subjects’ pupils were occluded due to blinking (in accordance with (Mathôt et al., 2013))). Artefacts in the pupil data were then smoothed using a moving average Hanning window (Kret & Sjak-Shie, 2019). Individual trials in which mean pupil

diameter while passively viewing the grey baseline screen was lower than 2mm or higher than 8mm were excluded ($N(\text{total trials from whole sample}) = 8$) as values outside this range are unnatural pupil sizes and were clear outliers based on inspection of participants' pupil-baseline histograms (Mathôt et al., 2018). Trials were averaged to form condition-specific pupil diameter waveforms to represent change in pupil size over time. Mean pupil diameter values during imagery in each trial were baseline corrected using a within-trial baseline subtraction approach (Mathôt et al., 2018) (i.e., subtracted from mean pupil diameter during 0.5 s prior to stimulus perception onset) to account for temporal shifts in pupil size across the experimental session due to fatigue (Morad et al., 2000). A two-way repeated measures analysis of variance (ANOVA) was used to compare Dark and Bright means during perception and imagery within both set-size conditions. 'Pupil-difference' scores were calculated by subtracting Dark condition means from Bright condition means of the corresponding set-size for comparison with binocular rivalry percent-primed scores. Pupil-difference scores were also separated based on the discrete within-trial vividness ratings to assess metacognition and whether pupil size changes in response to imagery were reflective of subjects' own experience of vividness of visual imagery.

In the binocular rivalry task, trials where participants reported seeing the pattern they were cued to imagine as dominant in the subsequent binocular rivalry display were denoted as 'primed' trials. The number of primed trials divided by the total number of trials (excluding mock trials and mixed percepts) was used to calculate a percent primed score for each participant. Participants' percent primed scores in binocular rivalry were correlated with their pupil-difference scores (both set-size-one and set-size-four) to assess potential for the pupillary response task to measure individual variability in visual imagery strength.

The LME's were run in R (Team, 2018) using the lme4 package and ANOVA's and ANCOVA were run in SPSS v.25.0 (IBM Corp. Released, 2017) and JASP was used to calculate effect sizes (JASP Team (2020). JASP Version 0.14.1). For the linear mixed-effects models set size (one or four) and vividness ratings (1, 2, 3 and 4) were entered into the model as fixed effects. As random effects intercepts for subjects were entered into the model. P-values were obtained by likelihood ratio tests of the full model with vividness included versus the model without vividness included.

Bayesian statistics were used to determine whether null findings can be interpreted as evidence for an absence of effect #ref. We used Bayesian repeated measure ANOVA (within-subject effect: set-size; between-subject effect: group) to compare the control and aphantasia groups as well as Bayesian one-sample t-tests to compare each group with H_0 , defined as the absence of effect. All Bayesian analyses were performed with JASP (Version 0.10.2).

Author Contributions: LK, RK and JP conceived of and designed the study. LK and RK recruited participants and collected the data. LK performed the pre-processing of the eye-pupil data under the supervision of TA and RK. LK, RK and TA all performed some of the data analysis. LK wrote the first draft of the paper with critical revisions from all authors. All authors approved the final version of the manuscript for submission.

References

- Akgul, F. (2013). *ZeroMQ : use ZeroMQ and learn how to apply different message patterns*. Packt Publishing.
- Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology : CB*, *23*(15), 1427–1431. <https://doi.org/10.1016/j.cub.2013.05.065>
- Bergmann, J., Genç, E., Kohler, A., Singer, W., & Pearson, J. (2016). Smaller primary visual cortex is associated with stronger, but less precise mental imagery. *Cerebral Cortex*. <https://doi.org/10.1093/cercor/bhv186>
- Binda, P., Pereverzeva, M., & Murray, S. O. (2013a). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.3440-12.2013>
- Binda, P., Pereverzeva, M., & Murray, S. O. (2013b). Pupil constrictions to photographs of the sun. *Journal of Vision*, *13*(6). <https://doi.org/10.1167/13.6.8>
- Blom, T., Mathôt, S., Olivers, C. N. L., & Van der Stigchel, S. (2016). The pupillary light response reflects encoding, but not maintenance, in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(11), 1716–1723. <https://doi.org/10.1037/xhp0000252>
- Bouffard, M. A. (2019). The Pupil. *CONTINUUM: Lifelong Learning in Neurology*, *25*(5), 1194–1214. <https://doi.org/10.1212/CON.0000000000000771>
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607. <https://doi.org/10.1111/j.1469-8986.2008.00654.x>
- Chang, S., & Pearson, J. (2018). The functional effects of prior motion imagery and motion perception. *Cortex*, *105*, 83–96. <https://doi.org/10.1016/j.cortex.2017.08.036>
- Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, *10*(1), 1–10. <https://doi.org/10.1038/s41598-020-65705-7>
- de Vito, S., & Bartolomeo, P. (2016). Refusing to imagine? On the possibility of psychogenic aphantasia. A commentary on Zeman et al. (2015). In *Cortex*. <https://doi.org/10.1016/j.cortex.2015.06.013>
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2017). Vividness of visual imagery depends on the neural overlap with perception in visual areas. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.3022-16.2016>
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2019). Shared Neural Mechanisms of Visual Perception and Imagery. In *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2019.02.004>
- Fotiou, D. F., Brozou, C. G., Tsiptsios, D. J., Fotiou, A., Kabitsi, A., Nakou, M., Giantselidis, C., & Goula, A. (2007). Effect of age on pupillary light reflex: Evaluation of pupil mobility for clinical practice and research. *Electromyography and Clinical Neurophysiology*.

- Francis Galton - “Statistics of mental imagery.” *Mind*(5 1880):301-18. (n.d.). Retrieved August 12, 2020, from http://galton.org/bib/JournalItem.aspx_action=view_id=102
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*. <https://doi.org/10.1016/j.cogbrainres.2004.02.012>
- Hustá, C., Dalmaijer, E., Belopolsky, A., & Mathôt, S. (2019). The pupillary light response reflects visual working memory content. *Journal of Experimental Psychology: Human Perception and Performance*. <https://doi.org/10.1037/xhp0000689>
- IBM Corp. Released. (2017). IBM SPSS Statistics version 25.0. In 2017.
- Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex*. <https://doi.org/10.1016/j.cortex.2017.10.014>
- Joshi, S., & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. In *Trends in Cognitive Sciences* (Vol. 24, Issue 6, pp. 466–480). Elsevier Ltd. <https://doi.org/10.1016/j.tics.2020.03.005>
- Just, M. A., Newman, S. D., Keller, T. A., McEleney, A., & Carpenter, P. A. (2004). Imagery in sentence comprehension: An fMRI study. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2003.08.042>
- 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>
- Kassner, M., Patera, W., & Bulling, A. (2014). Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. *UbiComp 2014 - Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. <https://doi.org/10.1145/2638728.2641695>
- Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PloS One*, 6(12), e29221. <https://doi.org/10.1371/journal.pone.0029221>
- Keogh, R., & Pearson, J. (2014a). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*. <https://doi.org/10.1167/14.12.7>
- Keogh, R., & Pearson, J. (2014b). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, 14(12). <https://doi.org/10.1167/14.12.7>
- Keogh, R., & Pearson, J. (2017). The perceptual and phenomenal capacity of mental imagery. *Cognition*. <https://doi.org/10.1016/j.cognition.2017.02.004>
- Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, 105, 53–60. <https://doi.org/10.1016/j.cortex.2017.10.012>
- 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>
- Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary stroop effects. *Cognitive Processing*. <https://doi.org/10.1007/s10339-010-0370-z>

- Laeng, B., & Sulutvedt, U. (2014). The Eye Pupil Adjusts to Imaginary Light. *Psychological Science*, 25(1), 188–197. <https://doi.org/10.1177/0956797613503556>
- Larsen, R. S., & Waters, J. (2018). Neuromodulatory correlates of pupil dilation. In *Frontiers in Neural Circuits*. <https://doi.org/10.3389/fncir.2018.00021>
- Lobato-Rincón, L. L., Cabanillas-Campos, M. del C., Bonnin-Arias, C., Chamorro-Gutiérrez, E., Murciano-Cespedosa, A., & Roda, C. S. R. (2014). Pupillary behavior in relation to wavelength and age. *Frontiers in Human Neuroscience*. <https://doi.org/10.3389/fnhum.2014.00221>
- Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, 1(1). <https://doi.org/10.5334/joc.18>
- Mathôt, S., Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V., & van der Sluis, S. (2013). A simple way to reconstruct pupil size during eye blinks. *Figshare*. <https://doi.org/10.1038/nn.3648>
- Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-017-1007-2>
- Mathôt, S., Grainger, J., & Strijkers, K. (2017). Pupillary Responses to Words That Convey a Sense of Brightness or Darkness. *Psychological Science*, 28(8), 1116–1124. <https://doi.org/10.1177/0956797617702699>
- Morad, Y., Lemberg, H., Yofe, N., & Dagan, Y. (2000). Pupillography as an objective indicator of fatigue. *Current Eye Research*, 21(1), 535–542. [https://doi.org/10.1076/0271-3683\(200007\)2111-ZFT535](https://doi.org/10.1076/0271-3683(200007)2111-ZFT535)
- Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage*, 105, 215–228. <https://doi.org/10.1016/j.neuroimage.2014.10.018>
- Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human Computer Studies*, 59(1–2), 185–198. [https://doi.org/10.1016/S1071-5819\(03\)00017-X](https://doi.org/10.1016/S1071-5819(03)00017-X)
- Pearson, J. (2014). New Directions in Mental-Imagery Research: The Binocular-Rivalry Technique and Decoding fMRI Patterns. *Current Directions in Psychological Science*. <https://doi.org/10.1177/0963721414532287>
- Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*. <https://doi.org/10.1038/s41583-019-0202-9>
- Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The Functional Impact of Mental Imagery on Conscious Perception. *Current Biology*. <https://doi.org/10.1016/j.cub.2008.05.048>
- Pearson, J., & Keogh, R. (2019). Redefining Visual Working Memory: A Cognitive-Strategy, Brain-Region Approach. *Current Directions in Psychological Science*. <https://doi.org/10.1177/0963721419835210>
- Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the mind’s eye: The metacognition of visual

- imagery. *Psychological Science*. <https://doi.org/10.1177/0956797611417134>
- Pounder, Z., Jacob, J., Jacobs, C., Loveday, C., Towell, T., & Silvanto, J. (2018). Mental rotation performance in aphantasia. *Journal of Vision*. <https://doi.org/10.1167/18.10.1123>
- Rademaker, R. L., & Pearson, J. (2012). Training visual imagery: Improvements of metacognition, but not imagery strength. *Frontiers in Psychology*, 3(JUL). <https://doi.org/10.3389/fpsyg.2012.00224>
- Sack, A. T., Camprodon, J. A., Pascual-Leone, A., & Goebel, R. (2005). Neuroscience: The dynamics of interhemispheric compensatory processes in mental imagery. *Science*. <https://doi.org/10.1126/science.1107784>
- Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The Future of Memory: Remembering, Imagining, and the Brain. In *Neuron*. <https://doi.org/10.1016/j.neuron.2012.11.001>
- Schacter, D. L., & Madore, K. P. (2016). Remembering the past and imagining the future: Identifying and enhancing the contribution of episodic memory. *Memory Studies*, 9(3), 245–255. <https://doi.org/10.1177/1750698016645230>
- Schwalm, M., & Jubal, E. R. (2017). Back to pupillometry: How cortical network state fluctuations tracked by pupil dynamics could explain neural signal variability in human cognitive neuroscience. *ENeuro*. <https://doi.org/10.1523/ENEURO.0293-16.2017>
- Shine, J. M., Keogh, R., O’Callaghan, C., Muller, A. J., Lewis, S. J. G., & Pearson, J. (2015). Imagine that: Elevated sensory strength of mental imagery in individuals with Parkinson’s disease and visual hallucinations. *Proceedings of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rspb.2014.2047>
- Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.0610082104>
- Team, R. C. (2018). R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2014. *Google Scholar*.
- Tsujimura, S., Wolffsohn, J. S., & Gilmartin, B. (2003). Pupil responses associated with coloured afterimages are mediated by the magno-cellular pathway. *Vision Research*, 43(13), 1423–1432. [https://doi.org/10.1016/S0042-6989\(03\)00145-7](https://doi.org/10.1016/S0042-6989(03)00145-7)
- van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. In *Psychonomic Bulletin and Review*. <https://doi.org/10.3758/s13423-018-1432-y>
- Wang, X., Ley, A., Koch, S., Hays, J., Holmqvist, K., & Alexa, M. (2020). Computational discrimination between natural images based on gaze during mental imagery. *Scientific Reports*, 10(1), 13035. <https://doi.org/10.1038/s41598-020-69807-0>
- Wassell, J., Rogers, S. L., Felmingam, K. L., Bryant, R. A., & Pearson, J. (2015). Sex hormones predict the sensory strength and vividness of mental imagery. *Biological Psychology*, 107, 61–68. <https://doi.org/10.1016/j.biopsycho.2015.02.003>

- Wicken, M., Keogh, R., & Pearson, J. (2021). The critical role of mental imagery in human emotion: insights from fear-based imagery and aphantasia. *Proc Biol Sci*, 288(1946), 20210267. doi:10.1098/rspb.2021.0267
- Xie, S., Kaiser, D., & Cichy, R. M. (2020). Visual Imagery and Perception Share Neural Representations in the Alpha Frequency Band. *Current Biology*, 30(13), 2621-2627.e5. <https://doi.org/10.1016/j.cub.2020.04.074>
- Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery - Congenital aphantasia. In *Cortex* (Vol. 73, pp. 378–380). Masson SpA. <https://doi.org/10.1016/j.cortex.2015.05.019>
- Zeman, A. Z. J., Della Sala, S., Torrens, L. A., Gountouna, V.-E., McGonigle, D. J., & Logie, R. H. (2010). Loss of imagery phenomenology with intact visuo-spatial task performance: a case of “blind imagination”. *Neuropsychologia*, 48(1), 145–155. <https://doi.org/10.1016/j.neuropsychologia.2009.08.024>
- Zokaei, N., Board, A. G., Manohar, S. G., & Nobre, A. C. (2019). Modulation of the pupillary response by the content of visual working memory. *Proceedings of the National Academy of Sciences*, 201909959. <https://doi.org/10.1073/pnas.1909959116>

Supplementary Information for:

The eyes have it: The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength

This file includes:

Figures S1 to S7

Legends for Figures S1 to S7

Table S1

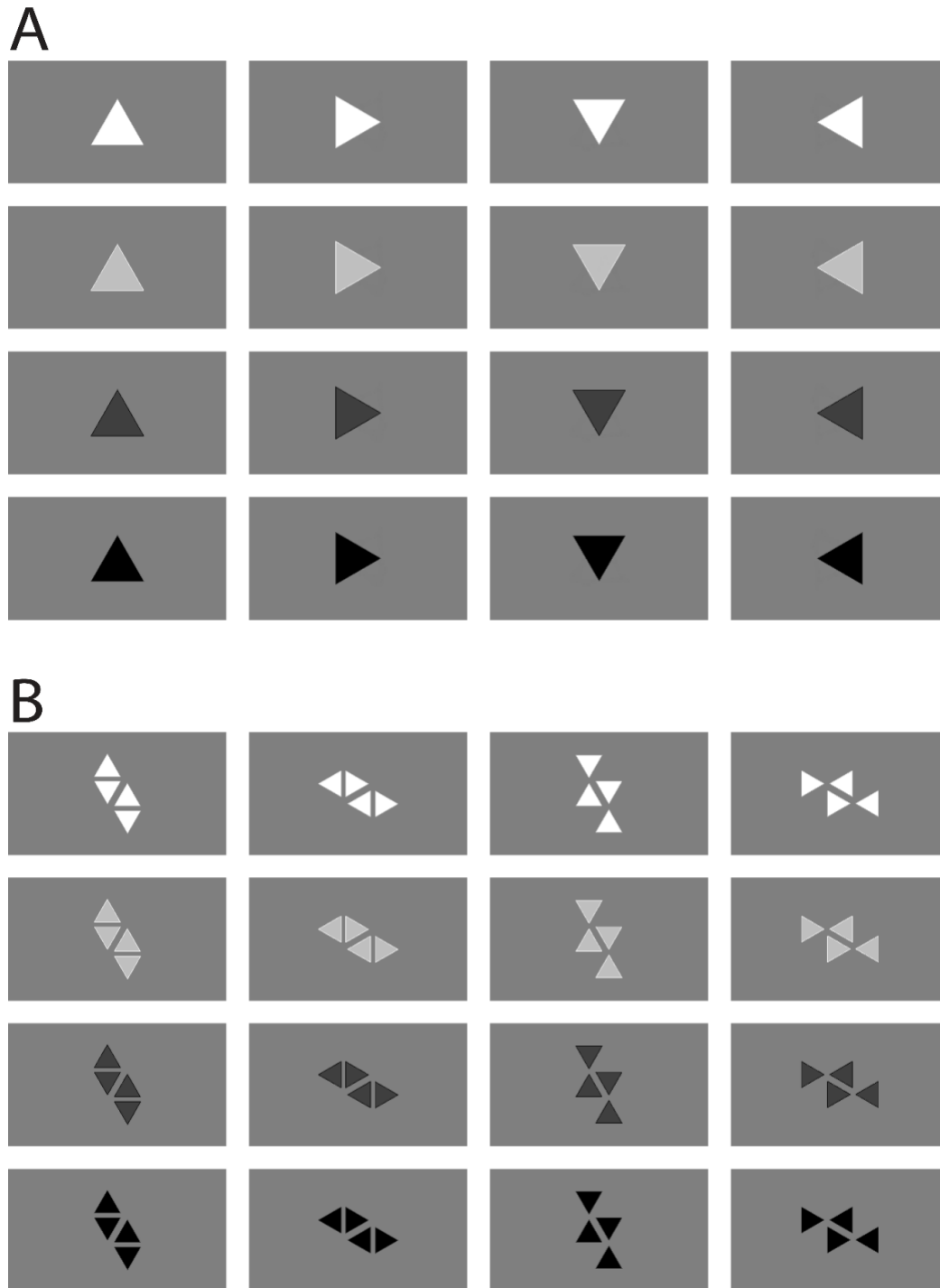


Fig. S1.

(A) Illustrations of all set-size-one stimuli used in the pupillary response task. (B) Illustrations of all set-size-four stimuli used in the pupillary response task.

Figure S2 shows the relationship between participants' mean trial-by-trial vividness ratings and their pupil-difference scores from set-size-one and set-size-four conditions, from the general population. Because assumptions of normality were met for the set-size-one data set, Pearson's correlational coefficient was used to assess the relationship between mean vividness ratings and set-size-one pupil-difference scores. No significant correlation was found between these two variables, $r_p(41) = -0.16$, $p = .31$ (see figure S2A). Due to the set-size-four data set violating normality, Spearman's correlational coefficient was used to analyse the relationship between set-size-four pupil-difference scores and VVIQ scores, with no significant correlation observed, $r_s(41) = .06$, $p = .69$ (see figure S2B).

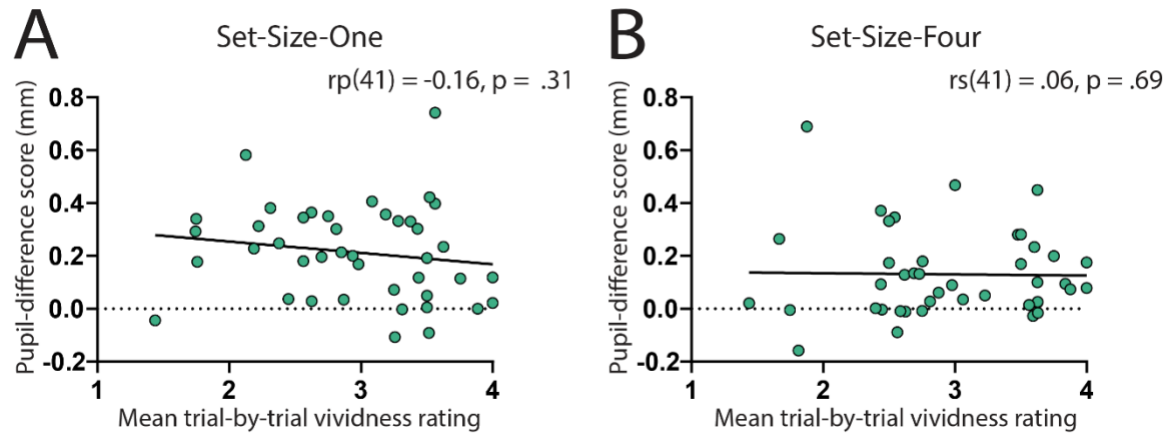


Fig. S2.

Correlation between visual imagery strength, as measured by the pupillary reflex task (pupil-difference score: difference between bright and dark conditions) and visual imagery strength as measured by mean trial-by-trial vividness ratings. (A) Set-size-one condition. (B) Set-size-four condition. Scatterplots show the general population (green circles and black trendline) data only. The horizontal dotted line indicates no effect of imagined luminance on pupil size. All data points represent one participant.

Figure S3 shows the relationship between participants' Vividness of Visual Imagery Questionnaire (VVIQ) scores and their pupil-difference scores from set-size-one and set-size-four conditions, from the general population. Again, assumptions of normality and sphericity were met for the set-size-one pupil data set but not the set-size-four pupil data set. On examining participants' Set-Size-One pupil-difference scores and their VVIQ scores, using the Pearson's correlational coefficient, no significant correlation was found, $r_p(41) = .02, p = .88$ (see figure S3A). Spearman's correlational coefficient indicated no significant correlation between set-size-four difference scores and mean vividness ratings, $r_s(41) = .20, p = .19$ (see figure S3B).

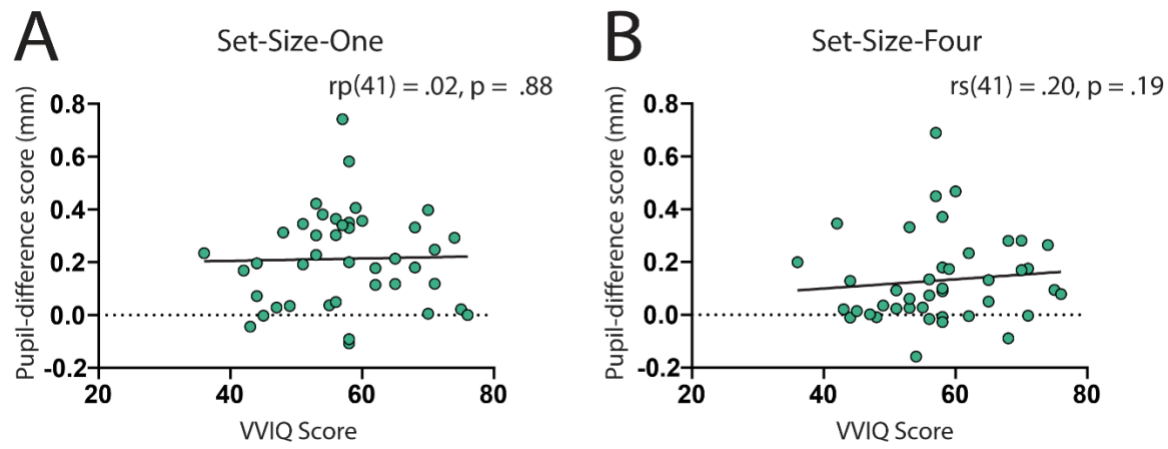


Fig. S3.

Correlation between visual imagery strength, as measured by the pupillary reflex task (pupil-difference score: difference between bright and dark conditions) and visual imagery strength as measured by the Vividness of Visual Imagery Questionnaire (VVIQ). (A) Set-size-one condition. (B) Set-size-four condition. Scatterplots show the general population (green circles and black trendline) data only. The horizontal dotted line indicates no effect of imagined luminance on pupil size. All data points represent one participant.

Figure S4 shows the mean vividness ratings for general population (left panel) and aphantasic individuals (right panels) for set size one and four for bright (red) and dark (blue) images. Aphantasic individuals score at floor on vividness ratings, for this reason we ran a repeated measures ANOVA only on the general population to assess what effect set size and image brightness had on vividness ratings. A 2 (Set Size: One and Four) x 2 (Stimuli: Bright and Dark) repeated measures ANOVA was run. There was a significant main effect of stimuli $F(1,41) = 37.85, p < .001$, with the general population rating imagery of dark images as being more vivid than bright images. The main effect of set size was also significant $F(1,41) = 8.042, p = .007$, with participants rating imagery for set size one as being more vivid than set size four. There was no significant interaction between the two factors $F(1,41) = 1.28, p = .264$.

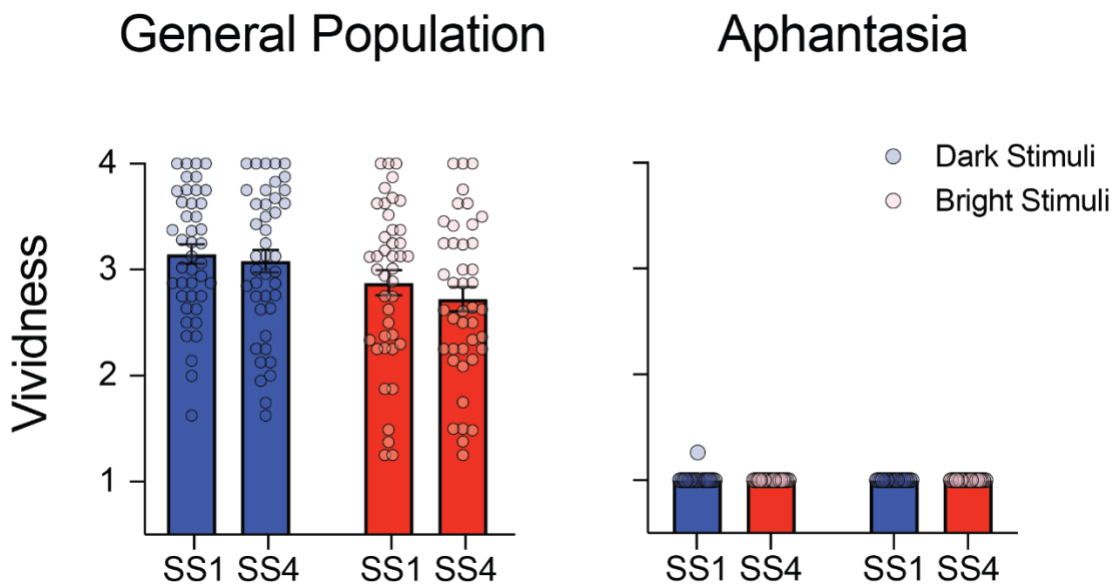


Fig. S4

Data shows the mean vividness ratings for general population (left panel) and aphantasic individuals (right panels) for set-size-one (SS1) and set-size-four (SS4) for bright (red) and dark (blue) images. Bars show mean values and error bars represent \pm SEM's. Each circle represents an individual.

Figure S5 shows the relationship between participants pupil size during perception and imagery. There are no significant correlations (all correlations are spearman's rho due to violation of normality for imagery SS4 difference score, imagery SS4 bright and dark).

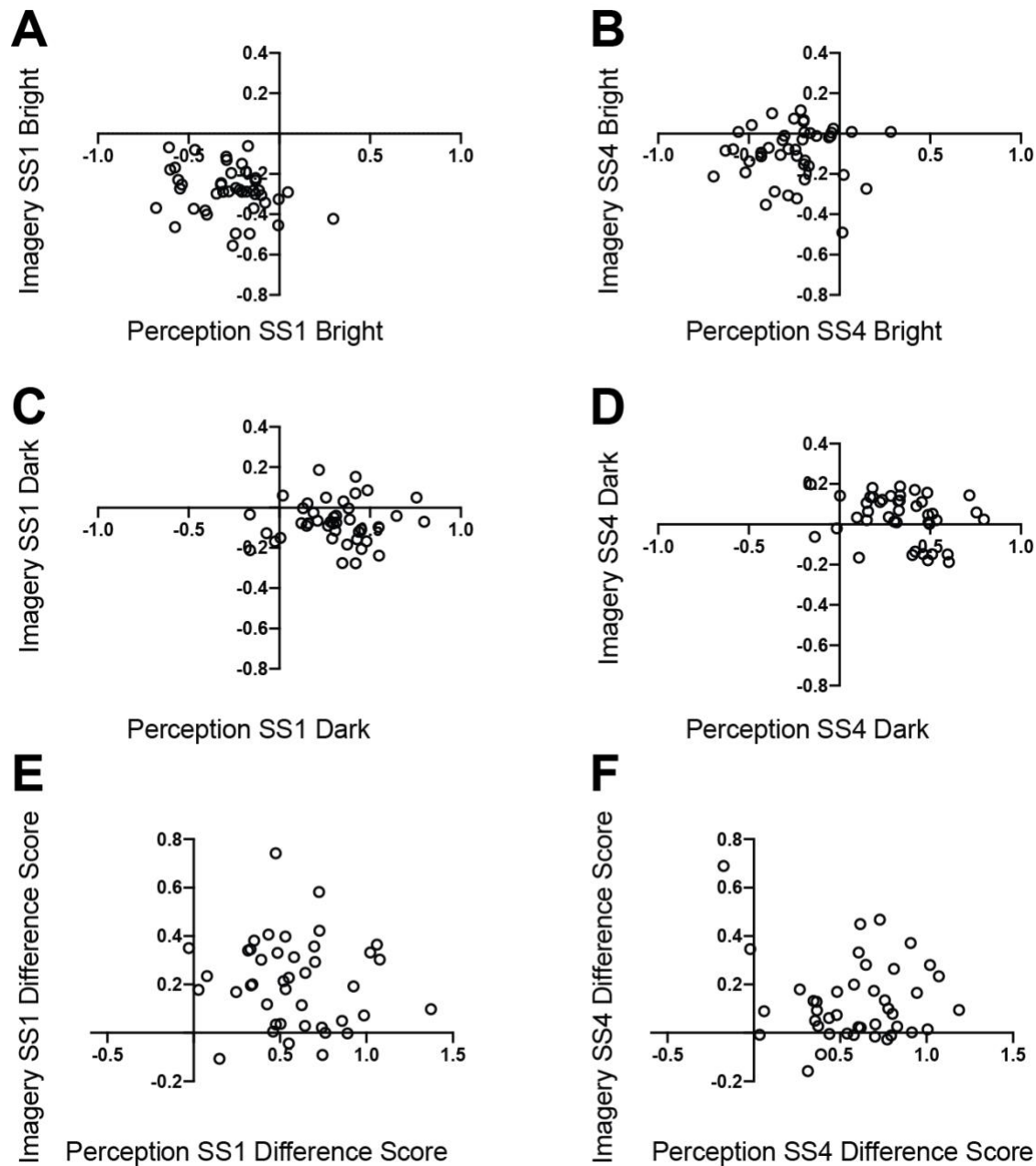


Fig. S5

Correlations between baselined pupil sizes during perceptually viewing images and imagining them for set size 1 bright images (**A**: $r_s = -.25$, $p = .11$), set size 4 bright images (**B**: $r_s = .12$, $p = .45$), set size 1 dark images (**C**: $r_s = -.05$, $p = .79$) and set size 4 dark images (**D**: $r_s = -.25$, $p = .11$), plus difference scores for set size 1 (**E**: $r_s = -.12$, $p = .44$) and set size 4 (**F**: $r_s = .07$, $p = .65$), all correlation values are Spearman's rho and were calculated in JASP.

Figure S6 shows the perceptual pupil responses to bright and dark images (for both set size 1 and 4) for control (left panel) and aphantasic (right panel) participants. A 2 (image: bright x dark) x 2 (set size (SS): 1 x 4) x 2 (group: aphantasic x controls) repeated measures ANOVA was run. There was a main effect of image $F(1,58) = 211.24, p < .001$ and set size $F(1,58) = 4.16, p = .046$. There was no significant effect of group $F(1,58) = 1.15, p = .288$ and there were no significant interactions (image x group: $F(1,58) = 1.57, p = .216$, SS x group: $F(1,58) = .06, p = .816$, image x SS: $F(1,58) = .01, p = .921$, image x SS x group: $F(1,58) = .03, p = .875$).

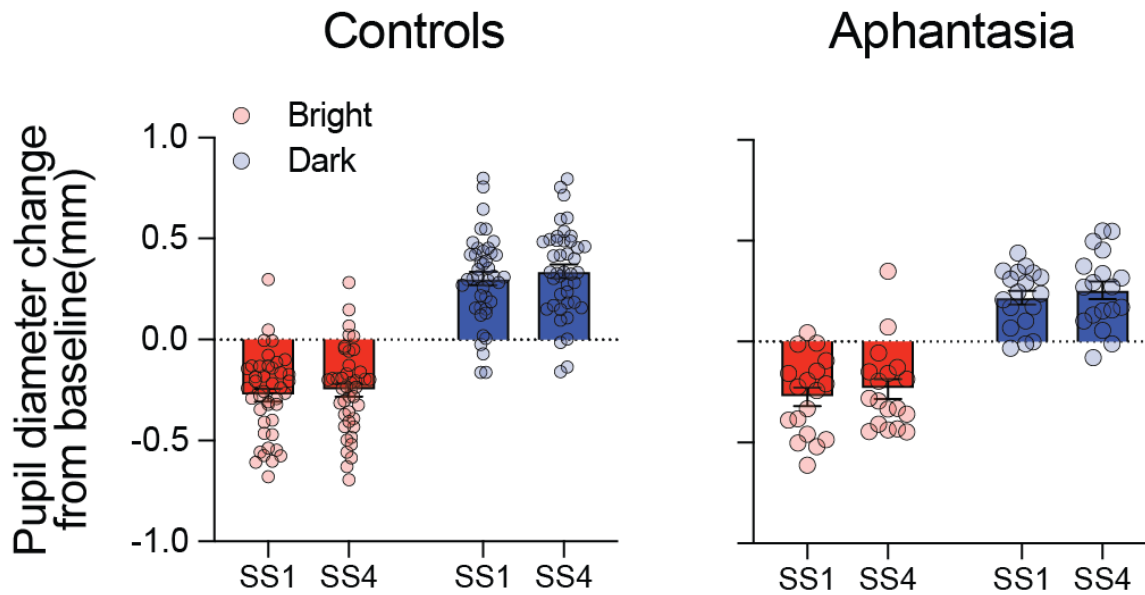


Fig. S6.

Data shows the perceptual pupil diameter change from baseline for control (left panel) and aphantasic individual's (right panel) for bright (red) and dark (images) for both set-size-one (SS1) and set-size-four (SS4). Bars show mean values and error bars represent \pm SEM's. Each circle represents an individual.

Figure S7 shows the relationship between participants pupil size difference scores during perception and priming in the binocular rivalry task. There are no significant correlations (all correlations are spearman's rho due to violation of normality).

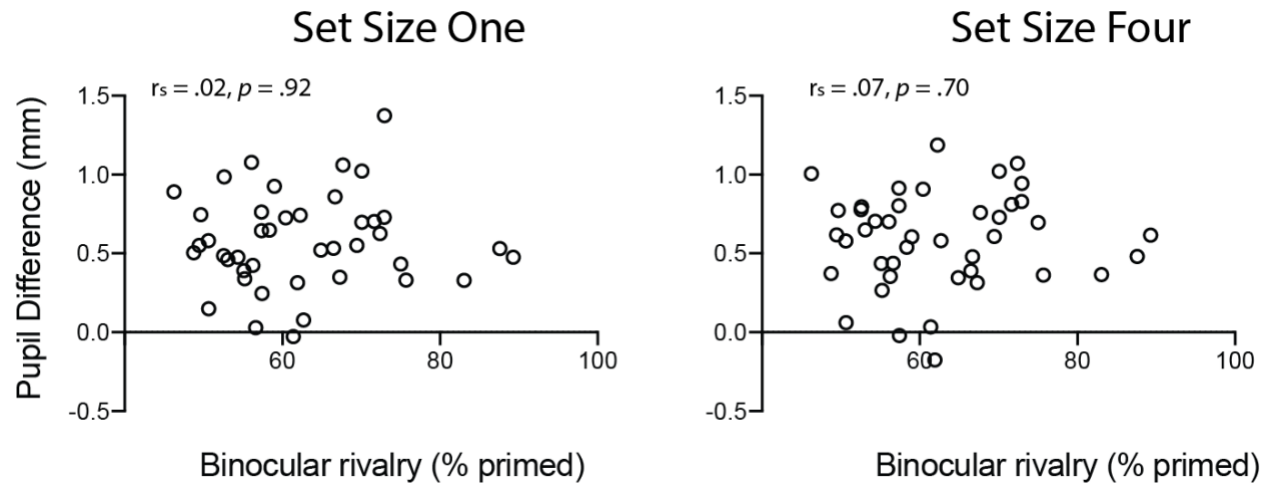


Fig. S7

Correlations between pupil difference scores when perceptually viewing images and binocular rivalry scores for set size one (**left panel:** $r_s = .02, p = .92$) and set size four (**right panel:** $r_s = .07, p = .70$).

Table S1

Fixed effects: Vividness Model (Pupil_dilation ~ Vivid_Rating + SetSize + (1 ID))			
	Estimate	Std. Error	t value
(Intercept)	0.15909	0.03645	4.365
Vivid_Rating2	0.07556	0.02563	2.948
Vivid_Rating3	0.17058	0.02488	6.855
Vivid_Rating4	0.14879	0.02528	5.887
SS	-0.07980	0.01663	-4.797

Table 1. Fixed effects for LME of vividness data (see Figure 1D)