Pupillary response to representations of light in paintings

Serena Castellotti — Department of Neurofarba, University of Florence, Florence, Italy

Martina Conti — Department of Neurofarba, University of Florence, Florence, Italy

Claudia Feitosa-Santana — Federal University of ABC, Sao Bernardo do Campo, Brazil; Neuroscience for Human Development, Sao Paulo, Brazil

Maria Michela Del Viva — Department of Neurofarba, University of Florence, Florence, Italy

It is known that, although the level of light is the primary determinant of pupil size, cognitive factors can also affect pupil diameter. It has been demonstrated that photographs of the sun produce pupil constriction independently of their luminance and other low-level features, suggesting that high-level visual processing may also modulate pupil response.

Here, we measure pupil response to artistic paintings of the sun, moon, or containing a uniform lighting, that, being mediated by the artist's interpretation of reality and his technical rendering, require an even higher level of interpretation compared with photographs. We also study how chromatic content and spatial layout affect the results by presenting grey-scale and inverted versions of each painting. Finally, we assess directly with a categorization test how subjective image interpretation affects pupil response.

We find that paintings with the sun elicit a smaller pupil size than paintings with the moon, or paintings containing no visible light source. The effect produced by sun paintings is reduced by disrupting contextual information, such as by removing color or manipulating the relations between paintings features that make more difficult to identify the source of light. Finally, and more importantly, pupil diameter changes according to observers' interpretation of the scene represented in the same stimulus.

In conclusion, results show that the subcortical pupillary response to light is modulated by subjective interpretation of luminous objects, suggesting the involvement of cortical systems in charge of cognitive processes, such as attention, object recognition, familiarity, memory, and imagination.

Introduction

The pupil is the central opening of the iris that regulates the intensity of light entering the eye to adjust retinal illumination and optimize vision (Loewenfeld, 1993). Light increments produce pupillary constriction (miosis), whereas light decrements produce pupillary dilation (mydriasis). This is known as pupillary light reflex (PLR), which is controlled by the autonomic nervous system (Gamlin & Clarke, 1995; Loewenfeld, 1993). Currently, a consistent body of evidence demonstrates that the PLR is not merely a basic low-level mechanism, showing that, even if the intensity of light is the primary determinant of the pupil size, non-visual factors can also affect the pupil diameter.

First studies of pupillometry showed that the pupil dilates not only in the dark but also in response to an increase in level of arousal (Bradley, Miccoli, Escrig, & Lang, 2008; Henderson, Bradley, & Lang, 2014; Hess & Polt, 1960; Snowden, O'farrell, Burley, Erichsen, Newton, & Gray, 2016), associated with an increased sympathetic activity (Aston-Jones & Cohen, 2005). Other studies demonstrated that the pupil dilates during the execution of mental tasks that require cognitive load (Beatty, 1982; Hess & Polt, 1964; Just & Carpenter, 1993), memory effort (Beatty & Kahneman, 1966; Goldinger & Papesh, 2012; Kafkas & Montaldi, 2011; Kahneman & Beatty, 1966; Papesh, Goldinger, & Hout, 2012), and decision making processes (De Gee, Knapen, & Donner, 2014; Einhäuser, Stout, Koch, & Carter, 2008).

More recent studies found that the pupil response can be modulated by high-level visual processes,

Citation: Castellotti, S., Conti, M., Feitosa-Santana, C., & Del Viva, M. M. (2020). Pupillary response to representations of light in paintings. *Journal of Vision*, *20*(10):14, 1–18, https://doi.org/10.1167/jov.20.10.14.

https://doi.org/10.1167/jov.20.10.14 Received May 18, 2020; published October 14, 2020 ISSN 1534-7362

such as attention (Binda & Murray 2015a; Binda, Pereverzeva, & Murray, 2013a; Binda, Pereverzeva, & Murray, 2014; Mathôt, van der Linden, Grainger, & Vitu, 2013; Naber, Alvarez, & Nakayama, 2013; Tkacz-Domb & Yeshurun, 2018; Unsworth, Robinson, & Miller, 2018), visual awareness in binocular rivalry conditions (Einhäuser et al., 2008; Fahle, Stemmler, & Spang, 2011; Kimura, Abe & Goryo, 2014; Naber, Frässle, & Einhäuser, 2011), perception of changes in stimuli low-level features, such as color or motion (Kohn & Clynes, 1969; Sahraie & Barbur, 1997; Ukai, 1985), perceptual illusions (Laeng & Endestad, 2012; Suzuki, Minami, Laeng, & Nakauchi, 2019; Zavagno, Tommasi, & Laeng, 2017), visual imagery (Laeng & Sulutvedt, 2014; Mathôt, Grainger, & Strijkers, 2017), and high-level processing of image content (Binda & Murray, 2015b; Binda, Pereverzeva, & Murray, 2013b; Naber & Nakayama, 2013; Sperandio, Bond & Binda, 2018).

Particularly relevant to the present study are findings showing that the pupil does not constrict only in response to the physical luminance of a stimulus, but also in response to its perceived luminance. For example, Lang and Endestad (2012) found that optical illusions that induce a subjective impression of brightness (Kitaoka lightness illusion) elicit pupillary constriction, compared with control stimuli (Kanizsa form illusion), despite the actual luminance was controlled. Later, Laeng and Sulutvedt (2014) continued the research toward an increasingly abstract level of stimuli, showing that mentally visualizing a bright scene, compared with a darker scene, produces pupillary constriction. Recently, Suzuki et al. (2019) found that colorful glare illusions (especially blue), that subjectively enhance the perception of brightness, induce pupillary constriction, reflecting an adaptive response of the visual system to a probable dangerous situation of dazzling sunlight. Furthermore, Binda et al. (2013b) found that pictures of the sun induce pupillary constriction compared with control stimuli of matched luminance, as photographs of the moon, showing that high-level interpretations of image content can modulate the pupil response. Naber and Nakayama (2013) also investigated the pupillary responses to a variety of natural scenes with the same low-level features, demonstrating a larger amplitude of pupil constriction to scenes containing a sun. By showing inverted images, they also investigated the effect of contextual information on the pupil, demonstrating how visual complexity affects pupil size. Taken together, these findings confirm that pupillary responses to ambient light reflect the interpretation of the light in the scene and not simply the amount of physical light energy entering the eye.

All of these studies indicate that the pupil diameter is sensitive to top-down modulation, and consequently that the pupil diameter could be modulated by cortical pathways other than the subcortical PLR system (Becket Ebitz & Moore, 2019; Binda & Murray, 2015a). A recent experiment (Sperandio et al., 2018) demonstrated that these extra-retinal modulations require visual awareness to modulate the pupil size. Using the continuous flash suppression (CFS) technique, they found that when participants were aware of sun pictures their pupils constricted relative to the control stimuli. This did not happen when the pictures were successfully suppressed from awareness, demonstrating that pupil size is sensitive to the contents of consciousness.

In the present study, we measured the pupil response to artistic paintings representing scenes with a visible sun, a visible moon, or the presence of diffused light to address the effect of cognitive interpretation of very complex stimuli. In fact, paintings render a scene through the artist's mind, requiring an even higher level of interpretation compared with photographs or artificial stimuli (Altschul, Jensen, & Terrace, 2017; Tatler & Melcher, 2007). In addition to the effect of image content, we also investigated the effect of contextual information, such as color and global layout. We aim to confirm that the pupil size depends on complex features of the visual stimulus that are presumably processed in the cortical areas.

The present study comprises one main and two control experiments to investigate effects of paintings categories, contextual information, and subjective interpretation (Figure 1).

Effects of paintings' categories

Three categories of paintings were used. Paintings of the sun were used to investigate if pictorial representations of high-luminance objects may elicit a smaller pupil size than other subjects, independently on the luminance of the images. Paintings of the moon were used to investigate pupillary response to stimuli representing a luminous disc as well but cognitively associated with a dark scene. Paintings with diffused light or different light sources (e.g. fires, volcanoes, etc.) were used to investigate if the mere presence of light in the absence of a luminous disc has any effect on pupil diameter.

To ensure that the results have general meaning, for each category we have purposely chosen stimuli painted over a period of more than 300 years and pertaining to very different styles, and we think this represents a strong point of the study.

In the main experiment (experiment 1), all the stimuli were presented by making them appear over a background of higher luminance. If the response depended only on overall light level, the same pupillary dilation would be expected for all stimuli. On the other hand, presentation of images depicting luminous objects is expected to produce pupil constriction due

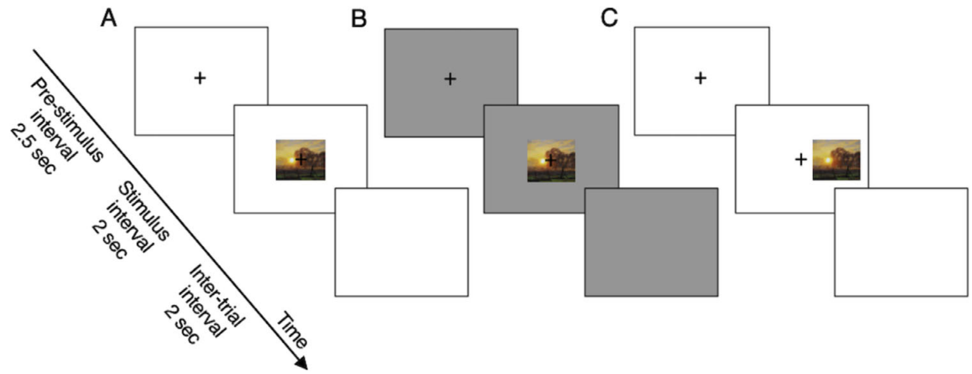

Figure 1. **Schematic of experimental procedures.** Procedure used in experiment 1 **(A)**, experiment 2 **(B)**, and experiment 3 **(C)**. Copyright permission from the Author was obtained for the painting shown, *Rural sunrise* (Gercken, 2012).

to high level visual processing (Binda et al., 2013b; Sperandio et al., 2018), over-riding the effect due to the physical properties of the stimulus. We expect to find a smaller pupil size for stimuli containing a light source, particularly the sun, due to the high-level interpretation of paintings content.

In a second control experiment (experiment 2), to rule out possible effects of luminance on the results of the main experiment, stimuli were presented by making them appear over a grey background of matching luminance. In this condition, there is no discrepancy between the luminance of the screen during fixation and the stimulus, therefore, any deviation from baseline pupil size would be due to stimulus content only. As in the main experiment, we expect a smaller pupil size for stimuli with luminous light sources.

Because studies have shown that pupillary responses are more sensitive to luminance changes in the fovea (Clarke, Zhang, & Gamlin, 2003), a third control experiment (experiment 3) was done by repeating the same paradigm of the main experiment, except stimuli were presented in the periphery of the visual field. We expect to confirm the results of the main experiment, ruling thus out as a possible dependence of pupillary response on retinal eccentricity.

Effects of contextual information

Color and spatial layout of images are crucial tools for artists to enhance the aesthetic experience in paintings (Chatterjee & Vartanian, 2014; Graham & Field, 2008; Montagner, Linhares, Vilarigues, & Nascimento, 2016; Nascimento et al., 2017). Color is a very important feature for interpreting visual scenes (Goffaux et al., 2004, Greene & Oliva 2009; Oliva & Schyns, 2000; Oliva & Torralba, 2006; Steeves et al., 2004); however, the effect of variations in stimulus color on pupil responses has been suggested, but not systematically investigated (Snowden et al., 2016). Contextual cues, such as relative position of objects and their orientation, are undoubtedly important for fast image interpretation (Oliva & Torralba, 2006). Disrupting these cues can have an effect in pupillary response to images, as already shown by Naber and Nakayama (2013) with computer rendering of natural images. These variables were investigated within experiment 1 and experiment 2, by comparing pupil responses to original paintings (up-right and full-color) with their inverted (180 degree rotated) and no-color (grey-scale) versions.

Effects of subjective interpretation

It is well known that aesthetic experience is unique to each individual (Kuchinke, Trapp, Jacobs, & Leder, 2009; Marković, 2010; Marković, 2011; Marković, 2012; Marković & Radonjić, 2008). It has also been shown that individual mental imagery (Laeng & Sulutvedt, 2014; Mathôt et al., 2017) and the content of consciousness (Sperandio et al., 2018) affect pupillary reactions. This means that the content represented in our paintings may be differently interpreted by each participant and, as a consequence, affect pupil diameter. For these reasons, we tested whether the paintings chosen as our stimuli elicited different pupil responses in experiment 1 depending on how the observer interpreted the scene, based on their response to a categorization test.

Methods

Participants

Twenty-eight observers (18 women and 10 men, mean age = 27.2, SD = 5) participated in experiment 1, another 12 observers (5 women and 7 men, mean age = 26.4, SD = 4) participated in experiment 2, and another 12 observers (6 women and 6 men, mean age = 26.5, SD = 4) participated in experiment 3. Before starting the experiments, all participants filled out a questionnaire about personal data, presence of aberration or optical defects, history of brain damage, medication intake, tobacco consumption, and caffeine intake. All selected participants had a normal or corrected-to-normal vision (by contact lenses) and did not take any type of medication. Participants were asked to abstain from drinking coffee before the experiment and not to wear eye make-up. Observers were unaware of the aim of the experiment and gave written informed consent before the experiment. All experimental procedures were approved by the local ethics committee (Comitato Etico Pediatrico Regionale – Azienda Ospedaliero-Universitaria Meyer – Firenze FI) and were compliant with the Declaration of Helsinki.

Apparatus and set-up

Each participant was tested individually in a dark room, with no lighting other than the display screen. Stimuli were presented on an ASUS monitor (51 × 29 cm, resolution 1920 × 1080 pixels), through a dedicated computer (iMac Retina 5K, 27-inch, mid 2015 3.3 GHz Intel Core i5 processor, MacOs Sierra software version 10.12.6). The observer was positioned at 57 cm distance from the monitor with a chin rest used to stabilize the head. Pupil diameter was binocularly tracked at 60 Hz with a CRS LiveTrack FM system (Cambridge Research Systems). Stimulus presentation and data collection programs were developed using Matlab (R2016b version).

Stimuli

We selected 30 paintings of natural scenes, produced in different historical periods (1700–2000) and with different styles (impressionism, realism, etc.). Each stimulus was nominally assigned to one of the three categories of our study, based on circumstantial elements, such as painting's title or the authors' interpretation (Table 1; for examples of each category see Figure 2A). All images were resized (conserving proportions) to either a width or a height of 283 pixels, with the other side ranging from 178 to 355 pixels. The original luminances of all paintings, in all their versions, were modified and were rescaled to the same value, corresponding to the average luminance of the whole set (9.7 cd/m^2). They were also rescaled to a common resolution (28.35 pixels/cm). The luminance varied within each image, reaching its maximum at the point where the source of illumination was represented. We measured the value of luminance at the center of each lunar/solar disc represented in our images, and tested for differences between sun and moon distributions, finding no statistically significant effect (sun: M = 40.2 cd/m^2, SD = 13.7 cd/m^2; moon: M = 37.5 cd/m^2, SD = 17.3 cd/m^2; $t(1) = 0.38$, $p > 0.05$; Figure 3).

In addition to the 30 paintings, a set of 10 uniform-grey rectangular images were generated, matching the mean luminance (9.7 cd/m^2) and the average size of paintings, to be used as control stimuli for luminance.

Furthermore, a grey-scale and an inverted (180 degree rotated) version were produced for each painting (see Figure 4A). They were used in experiments 1 and 2, to assess the role of color and global image organization.

Procedure

The eye tracker was calibrated at the beginning of each session with a standard 9-point calibration routine. In experiment 1, trials started with the presentation of a black fixation cross (5 × 5 mm) in the center of a white screen (71 cd/m^2) for 2.5 seconds (pre-stimulus interval). This was followed by the presentation of one of the stimuli for 2 seconds (stimulus interval). The fixation cross was kept visible in the center of the screen during the *pre-stimulus* and *stimulus intervals,* whereas the luminance of the background screen was kept constant at 71 cd/m^2. Observers were instructed to keep their gaze at the fixation cross for the whole of the pre-stimulus and stimulus intervals, refraining from blinking, and not to perform any other task. During this time, pupil size was continuously monitored by means of a camera attended by the experimenter on her own screen (using QuickTime software) throughout the whole experiment. Each trial was followed by an inter-trial interval of 2 seconds, in which a white screen (71 cd/m^2) was displayed. During this time, the eye tracker did not record, and the observers were allowed to blink and rest their eyes before the next trial (Figure 1A).

Experiment 1 consisted of 100 trials divided into 4 blocks of 25 images: 10 different paintings per category plus their inverted and grey-scale versions, plus 10 uniform-grey control stimuli. The sequence of stimuli presentation was randomly predetermined and kept the same for all observers.

Experiment 2 followed the same procedure of experiment 1, except that stimuli were presented on a grey background having the same luminance as the mean luminance of the stimuli (9.7 cd/m^2; Figure 1B). In this experiment, uniform grey control stimuli, having the same luminance as the background, were not used. This led to 90 trials in 2 blocks of 22 plus 2 blocks of 23 stimuli.

Paintings category	n°	Artist	Title	Year
Paintings of the sun	1	Loren D. Adams	Golden Sunset Reef	2012
	2	Graham Gercken	Rural sunrise	2012
	3	G. Peine Toomalatai	Precious sight	2009
	4	Albert Bierstadt	Aurora	1850
	5	Vincent Van Gogh	The sower	1888
	6	Abraham Hunter	Evening mist	2000
	7	Debbie Cusick	St. Johns sunrise	2012
	8	Ken Bushe	Tentsmuir beach	2000
	9	Frederic Edwin Church	The andes of Ecuador	1876
	10	Z. L. Feng	Watercolor landscape	2000
Paintings of the moon	11	Phyllis Gates	Full moon on the Pacific	2018
	12	Donato Creti	Osservazione astronomica della luna	1711
	13	Katie Larner	Silver moon	2012
	14	Massimo Cavallari	La luna del cacciatore	2005
	15	Bruno Lucatello	Notte di luna veneziana	2000
	16	René Magritte	Le Maître d'école	1955
	17	Vincent van Gogh	Starry night	1889
	18	Barbara Solberg	Harvest moon	2012
	19	Mesheryakov	Caribbean night ocean	2000
	20	Lovell Birge Harrison	Moonlight over a pond	1900
Paintings with diffused light (or other sources of illumination)	21	Claude Monet	Landscape at Giverny	1888
	22	Laurent Parcelier	Gardens	1996
	23	William Turner	The slave ship	1840
	24	Hans Dahl	Upon sunny waves	1900
	25	William Turner	Fort vimieux	1831
	26	Paul Dougherty	Waves crashing on the rocks	1900
	27	Robert Finale	Costa Azul	2006
	28	William Turner	Eruption of Vesuvius	1817
	29	Claude Monet	Haystacks at Giverny, the evening sun	1888
	30	William Turner	The Burning of the Houses of Lords and Commons	1835

Table 1. List of paintings used as stimuli.

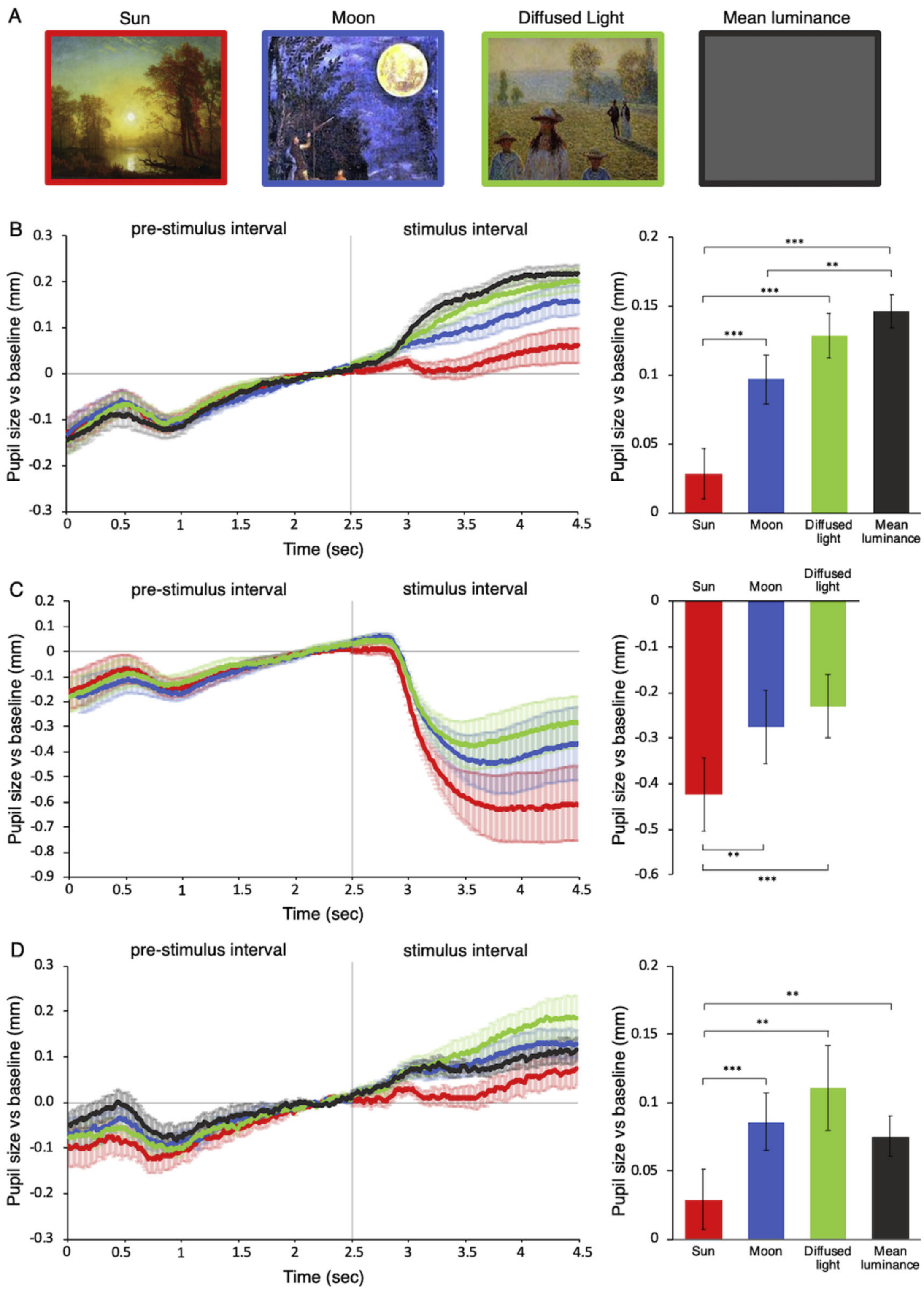

Figure 2. **Mean pupillary responses to different paintings categories. (A)** Examples of paintings belonging to the four categories of stimuli. Example of sun painting: *Aurora* (Bierstadt, 1850); example of moon painting: *Astronomical Observations: the Moon* (Creti, 1711); example of diffused light painting: *Landscape at Giverny* (Monet, 1888). Paintings shown are in the public domain. **(B)** Experiment 1. **Left:** Baseline-corrected pupil size $\overline{p}(t)$, for the four stimulus categories, plotted as a function of time from trial onset. **Right:** μ of different categories. **(C)** Experiment 2. **Left:** $\overline{p}(t)$for the three stimulus categories. **Right:** μ of different categories. **(D)** Experiment 3. **Left:** $\bar{p}(t)$ for the four stimulus categories. **Right:** μ of different categories. The vertical line in the graphs on the left indicates stimulus onset. Error bars on the left are $SE(t)$. Error bars on the right are SE of the means μ. Red: sun; blue: moon; green: diffused light; black: grey-uniform control stimuli. Asterisks mark statistically significant pairwise comparisons across image categories: $*p < 0.05$; $**p < 0.01$, $***p < 0.001$. All data shown have been corrected based on each observer's categorization.

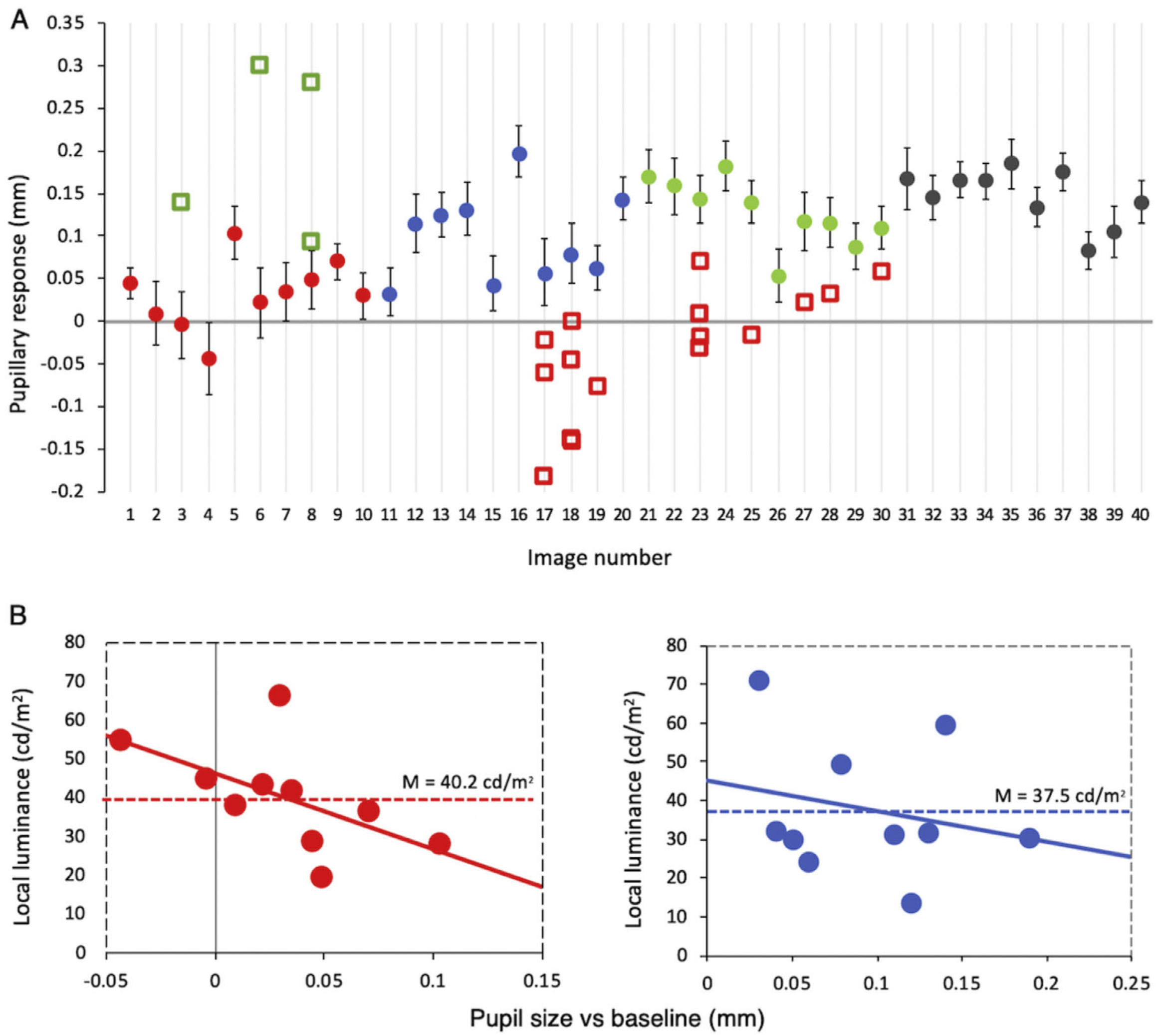

Figure 3. **Pupillary responses to individual stimuli. (A)** Round filled symbols are average responses for each image, μ_i, of observers that classified the paintings according to the nominal classification given by the authors. Red: sun paintings; blue: moon paintings; green: diffused light paintings and grey: uniform-grey control stimuli. Hollow squares are individual responses of observers that did not classify the paintings according to the nominal classification. Red: painting classified as *sun*, green: painting classified as *other*. Error bars are SE_i. Locations of misinterpretations in the distribution of each image are Image 3: 82nd percentile; image 6: 96th percentile; image 8: 57th < percentile < 93rd; image 17: 11th < percentile < 39th; image 18:11th < percentile < 46th; image 19: 21st percentile; image 23: 11th < percentile < 39rd; image 25: 18th percentile; image 27: 39th percentile; image 28: 21st percentile; image 30: 32nd percentile. **(B)** Correlation between the local luminance at the center of the light source of each painting and the corresponding pupillary response averaged across observers μ_i (experiment 1). There is no significant correlation between pupil dilation and local luminance at the center of suns ($R^2 = 0.23$, $F(1) = 3.83$, $p > 0.05$) or moons ($R^2 = 0.06$, $F(1) = 0.45$, $p > 0.05$). The dotted lines indicate the mean luminance in the center of sun (red, M = 40.2 cd/m^2, SD = 13.7 cd/m^2) and moon paintings (blue, M = 37.5 cd/m^2, SD = 17.3 cd/m^2). All data shown have been corrected based on each observer's categorization.

Experiment 3 also followed the same procedure of experiment 1, but stimuli were presented in an off-center location, 5 degrees to the right of the fixation cross (Figure 1C). In this case, grey-scale and inverted versions of paintings were not tested, leading to 40 trials divided in 2 blocks of 20 stimuli.

After the experiments, all paintings were presented again in sequence to the observers without time limitation and pupil recording, asking them to categorize each, as "sun," "moon," or "other." The complete procedure took about 50 minutes per observer, of which about 30 minutes were of pupil recordings.

Data processing

Raw data recorded by the eye-tracker were processed in the same way for all three experiments. Right and left

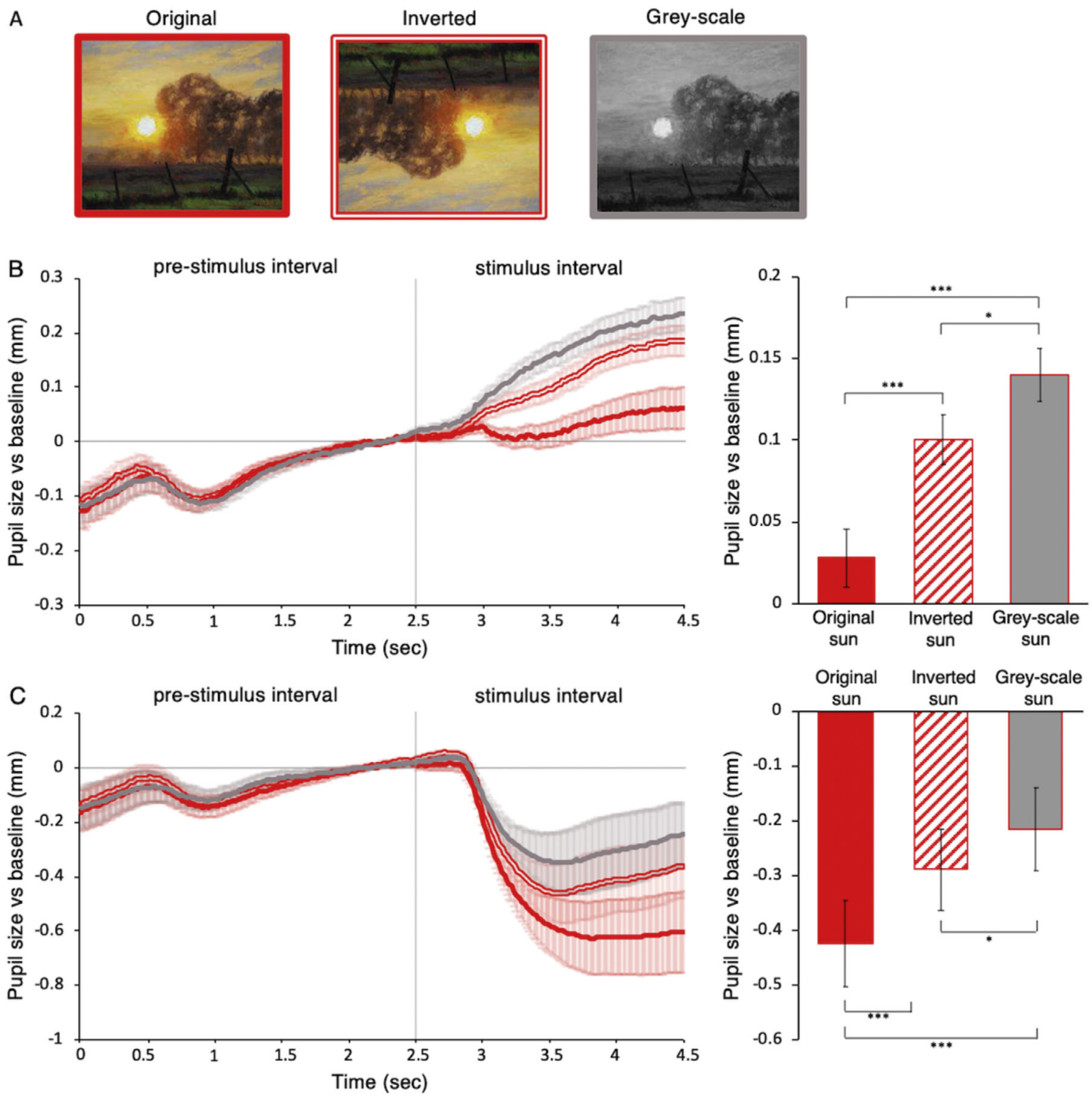

Figure 4. **Mean pupillary responses to different versions of paintings of the sun. (A)** Example of a sun painting in original, inverted and grey-scale version. Copyright permission from the author was obtained for the painting shown, *Rural sunrise* (Gercken, 2012). **(B)** Experiment 1. **Left:** Baseline-corrected pupil size $\overline{p}(t)$, for the three versions of sun paintings, plotted as a function of time from trial onset. **Right:** μ of different versions. **(C)** Experiment 2. **Left:** $\overline{p}(t)$ for the three versions of sun paintings. **Right:** μ of different versions. The vertical line in the graphs on the left indicates stimulus onset. Error bars on the left are the $SE(t)$. Error bars on the right are the SE of the means μ. Red: original versions of sun paintings; red/white: inverted versions of sun paintings; grey: grey-scale versions of sun paintings. Asterisks mark statistically significant pairwise comparisons across image categories: *$p < 0.05$; **$p < 0.01$, ***$p < 0.001$. All data shown have been corrected based on each observer's categorization.

pupil diameters were averaged, and the resulting value was transformed from pixels to millimeters. Calibration was attained by measuring the instrument's recording of a 4 mm artificial pupil, positioned at the approximate location of the subjects' left eye.

For each observer, a baseline pupil diameter was calculated by averaging pupil diameter recorded over the last 500 ms of the pre-stimulus interval in each trial. This baseline value was then subtracted from each recording of that observer over the whole 4.5 second period (Mathôt, Fabius, Van Heusden, & Van der Stigchel, 2018).

All results were classified according to the categorization made by the observer in the test, to ensure that the pupil size corresponded to the subjective interpretation of the nature of light source. For example, if a painting with a moonlit scene had been categorized as "sun" by some participants, the recordings obtained with this image were analyzed as a sun stimulus for this observer.

The analysis of the pupil responses elicited by different categories of paintings, or different versions of the same painting, follows a method widely used in literature for this type of experiments (Binda et al., 2013b; Naber & Nakayama, 2013). An average pupil size μ was calculated for each image category as follows. First, all recordings from each observer s $p_{s,i}(t)$, where i is the stimulus index, were averaged as a function of time $\overline{p_s}(t) = \frac{\sum_i^I p_{s,j}(t)}{I}$ (I = 10 for each category). Then, temporal averages were computed over the duration of the stimulus interval for each observer $\mu_s = \frac{\sum_{t=0}^{T} \overline{p_s}(t)}{T}$, from which the overall average was computed for each category as: $\mu = \frac{\sum_{s=1}^{N} \mu_s}{N}$. This quantity was attributed an overall variance SE $\sigma^2 = \frac{\sum_{s=1}^{N} (\mu_s - \mu)^2}{N-1}$. Differences between μ of different categories were assessed with ANOVA and pairwise comparisons were done with post hoc Student's t-tests with Bonferroni corrections.

Moreover, the response functions for each category $\overline{p}_s(t)$were averaged over the N observers, to obtain time-dependent averages $\overline{p}(t) = \frac{\sum_s^N \overline{p}_s(t)}{N}$ together with their time-dependent standard errors $SE(t) = \sqrt{\frac{\sigma^2(t)}{N}} = \sqrt{\frac{\sum_{s=1}^{N} (\overline{p_s}(t) - \overline{p}(t))^2}{N(N-1)}}$.

In addition, data were also analyzed on an image by image basis as follows. For each image i, the response for each participant s as a function of time, $p_{s,i}(t)$, was averaged over the *stimulus interval* to yield $\mu_{s,i} = \frac{\sum_{t=0}^{T} p_{s,i}(t)}{T}$ and then over all participants to yield the time-average response for each image $\mu_i \frac{\sum_{s=1}^{N} \mu_{s,i}}{N}$, with an associated standard error $SE_i = \sqrt{\frac{\sum_{s=1}^{N} (\mu_{s,i} - \mu_i)^2}{N\sqrt{N-1}}}$.

Results

Effects of paintings' categories

The main result of this work comes from the comparison of responses to the presentation of the three categories of paintings and to the uniform grey control stimuli. The time course of pupil size for each painting category $\bar{p}(t)$ obtained from experiment 1 is shown in Figure 2B (left). Because all images equally and greatly reduce the luminance level across the screen, if the response were based only on luminance, we would expect the same pupillary dilation for all categories. In fact, the line graph in Figure 2B (left) shows that sun stimuli elicited a much smaller dilation than all other categories, despite having the same mean luminance. Paintings with the moon, paintings with diffused light, and uniform grey control stimuli induced a consistent pupillary dilation.

Significant differences between all categories of stimuli μ are evidenced by ANOVA ($F(3) = 20.54$, $p < 0.001$). Pairwise comparisons (Table 2) show that paintings with the sun produced lower dilation than paintings with the moon, with diffused light and uniform luminance images. In addition, moon paintings produce smaller dilation than uniform grey control stimuli. No statistical difference is found between the dilation induced by diffused light paintings and moon or uniform grey control stimuli (see Figure 2B, right).

The size of differences between conditions, estimated by Cohen's d statistics, is *very small* for sun versus moon paintings ($s = 0.55$, $d = 0.12$), *small* for sun versus diffused light paintings ($s = 0.55$, $d = 0.18$), and sun versus uniform-grey ($s = 0.53$, $d = 0.22$), and *very small* for moon versus uniform grey ($s = 0.53$, $d = 0.09$). Values lower than 0.01 were considered to be negligible effects (Cohen, 1988; Savilowsky, 2009).

The time course $\overline{p}(t)$ also show the same general trend for all categories (see Figure 2B, left). Pupil diameter increases gradually during the pre-stimulus interval, then remains stable for about 500 ms after stimulus onset, at a common level for all categories. After this, pupil size starts to increase with different slopes according to different stimulus categories. The associated uncertainty $SE(t)$ also increases with time for painting stimuli, while staying approximately constant for control stimuli (see the Discussion section for possible explanations). This highlights the advantage pertaining to the second method of analysis, whereby different data points are combined with proper accounting for their differing uncertainties.

Because eye movements can influence pupil changes (Gagl, Hawelka, & Hutzler, 2011), although observers were instructed to keep fixation and their eye movements were monitored, we analyzed a posteriori the average position of their eyes with respect to the fixation cross for the different stimulus categories. The average distance from fixation in millimeters was minimal (sun: 2.45 ± 0.5; moon: 2.88 ± 0.6; diffused light: 2.09 ± 0.4; and mean luminance: 2.59 ± 0.5) and the same for all categories, included the uniform grey stimuli (ANOVA, $F(3) = 0.38$, $p > 0.05$).

Results of experiment 2 are displayed in Figure 2C. In this case, the same pupillary constriction is expected for all kinds of paintings, but we found that the constriction induced by paintings of the sun is larger than those elicited by paintings of the moon and paintings with diffused light (ANOVA ($F(2) = 11.88$, $p < 0.001$; see Table 2). The size of this effect is categorized as small for sun versus moon paintings ($s = 0.71$, $d = 0.2$) and sun versus diffused light ($s = 0.69$, $d = 0.3$).

In experiment 3, where paintings are displayed in the periphery of the visual field, the time course of responses (Figure 2D, left) suggests a lower dilation for paintings of the sun than for other categories. This

				Pairwise comparisons of µ *t*-tests (Bonferroni correction)		
	Painting category	M	SE	Moon	Diffused light	Mean luminance
Experiment 1	**Sun**	0.03	0.02	t (3) = 4.22 ***p* < 0.001*****	t (3) = 5.94 ***p* < 0.001*****	t (3) = 6.41 ***p* < 0.001*****
	Moon	0.10	0.02		t (3) = 2.05 *p* = 0.29	t (3) = 3.01 ***p* < 0.05***
	Diffused light	0.13	0.02	t (3) = 2.05 *p* = 0.29		t (3) = 1.37 *p* = 1
	Mean luminance	0.15	0.01	t (3) = 3.01 ***p* < 0.05***	t (3) = 1.37 *p* = 1	
Experiment 2	**Sun**	−0.42	0.08	t (2) = 3.18 ***p* < 0.05***	t (2) = 7.87 ***p* < 0.001*****	
	Moon	−0.28	0.08		t (2) = 0.93 *p* = 1	
	Diffused light	−0.23	0.07	t (2) = 0.93 *p* = 1		
Experiment 3	**Sun**	0.03	0.02	t (3) = 5.51 ***p* < 0.001*****	t (3) = 4.88 ***p* < 0.01****	t (3) = 4.06 ***p* < 0.01****
	Moon	0.09	0.02		t (3) = 1.21 *p* = 1	t (3) = 1.00 *p* = 1
	Diffused light	0.11	0.03	t (3) = 1.21 *p* = 1		t (3) = 1.71 *p* = 0.68
	Mean luminance	0.08	0.01	t (3) = 1.00 *p* = 1	t (3) = 1.71 *p* = 0.68	

Table 2. Statistics tests for effects of paintings categories.

is confirmed by the ANOVA analysis ($F(3)$ =9.86, $p <$ 0.001; see Table 2; Figure 2D, right). The size of this effect is very small for sun versus moon paintings ($s =$ 0.51, $d = 0.1$), sun versus grey uniform ($s = 0.49$, $d =$ 0.1), and small for sun versus diffused light ($s = 0.54$, $d = 0.2$).

Image by image analysis

Paintings are less uniform stimuli than photographs in representing a given subject. To assess the variance of the responses elicited by different paintings, data of experiment 1 have been analyzed image by image, and results are shown in Figure 3A.

The first finding is that the large majority of images were classified by observers in agreement with the nominal classification provided by the authors, but there were a small number of exceptions. They occur in 11 paintings, for a total of 20 observations, amounting to 2% of total occurrences. They are an interesting effect that we investigate further in the section Effects of subjective interpretation below, but their limited number has a small effect on the overall results, as we verified by repeating the analysis based on the nominal rather than the observers' classification.

For the cases where the paintings were perceived according to their nominal categorization, the variances in pupil responses ($\sigma^2_{SUN} = 0.002$, $\sigma^2_{MOON} = 0.003$, $\sigma^2_{DIFFUSED} = 0.002$, $\sigma^2_{GREY} = 0.001$) are compatible among all stimulus categories (Fisher's tests, $p >$ 0.1 for all comparisons). More importantly, they are also statistically compatible with the variance of the responses observed to uniform grey control stimuli (Fisher's tests, $p > 0.1$ for all comparisons). This indicates that the obvious differences between individual paintings do not dominate the observed spread in response.

For the cases where the paintings were not perceived according to their nominal categorization, pupil responses were always in the direction of the average of the perceived stimulus: when sun paintings, were perceived as other, pupil sizes were larger, when moon and diffused light paintings were perceived as sun, pupil sizes were smaller. However, values, although apparently off-scale, were all comprised within 11th and the 93rd percentiles of image distributions (for all values, see the caption of Figure 3A).

All stimuli had the same mean luminance, but they depict light sources of different size and intensity. To control for dependence on these variables, measurements in experiment 1 were correlated with the luminance value in the center of the light

				Pairwise comparisons of μ *t*-tests (Bonferroni correction)	
	Painting category	M	SE	Inverted sun	Grey-scale sun
Experiment 1	**Original sun**	0.03	0.02	t (2) = 4.72 **$p < 0.001$*****	t (2) = 7.40 **$p < 0.001$*****
	Inverted sun	0.10	0.02		t (2) = 2.65 **$p < 0.05$***
	Grey-scale sun	0.14	0.02	t (2) = 2.65 **$p < 0.05$***	
Experiment 2	**Original sun**	−0.42	0.08	t (2) = 7.28 **$p < 0.001$*****	t (2) = 6.38 **$p < 0.001$*****
	Inverted sun	−0.29	0.07		t (2) = 3.01 **$p < 0.05$***
	Grey-scale sun	−0.22	0.08	t (2) = 3.01 **$p < 0.05$***	

Table 3. Statistics tests for effects of contextual information.

source. Figure 3B shows no significant correlation between pupil dilation and local luminance at the center of suns ($R^2 = 0.23$, $F(1) = 3.83$, $p > 0.05$) or moons ($R^2 = 0.06$, $F(1) = 0.45$, $p > 0.05$). In addition, no statistical difference is seen between average local luminance values at the centers of the sun and moon light sources ($t(1) = 0.38$, $p > 0.05$).

Effects of contextual information

Another interesting result of experiment 1 follows from the comparison between pupillary response elicited by paintings of the sun in their original, grey-scale, and inverted versions (examples in Figure 4A). The graph Figure 4B (left) shows the time course of pupil size $\overline{p}(t)$ for sun paintings, and their grey-scale and inverted versions. Average pupil responses μ are found to be different between these three conditions (ANOVA: $F(2) = 28.09$, $p < 0.001$). Grey-scale and inverted versions produce a significantly wider pupillary dilation than the original version of the sun paintings. This suggests that manipulations of image structure or color may alter the interpretation of scene brightness and, as a consequence, modulate the pupil response itself. In addition, grey-scale versions produce a larger dilation than inverted versions of the paintings. This indicates that the global arrangement of painted elements is less important than their color in suggesting the presence of light in a painting (Table 3; Figure 4B, right). The size of these differences, assessed by Cohen's *d*, is very small for original versus inverted versions ($s = 0.54$, $d = 0.13$) and inverted versus grey-scale ($s = 0.54$, $d = 0.08$), and small for original versus grey-scale versions ($s = 0.55$, $d = 0.21$). ANOVA shows statistical differences also for different versions of diffused light paintings (ANOVA: $F(2) = 5.10$, $p < 0.01$). Indeed, grey-scale versions of diffused light paintings produce more dilation than their original versions ($t(2) = 3.04$, $p < 0.05$). Instead, responses to different versions of moon paintings are not statistically different (ANOVA: $F(2) = 1.87$, $p > 0.05$).

Although the same observer sees the same painting only once in the original, once in the reversed, and once in the grey-scale version, that are different for contextual information, there still may be a habituation effect on pupil size as described by Yoshimoto, Imai, Kashino, and Takeuchi (2014). A 2-way ANOVA ruled out this possibility showing a significant main effect of sun paintings' versions (ANOVA: $F(2) = 28$, $p < 0.001$) but no significant effect of order presentation ($F(2) = 1.28$, $p > 0.05$).

The same pattern of results is obtained with the same stimuli in experiment 2 (see Figure 4C, left). Original versions of sun paintings elicit more constriction than their inverted versions, that in turn elicit more constriction than grey scale versions (ANOVA: $F(2) = 33.14$, $p < 0.001$; see Table 3; Figure 4C, right). The size of these differences, assessed by Cohen's d, is small for original versus inverted versions ($s = 0.70$, $d = 0.2$) and original versus grey-scale ($s = 0.70$, $d = 0.3$), and very small for inverted versus grey-scale versions ($s = 0.70$, $d = 0.11$). ANOVA shows statistical differences also for different versions of moon (ANOVA: $F(2) = 5.96$, $p < 0.01$) and diffused light paintings (ANOVA: $F(2) = 15.48$, $p < 0.001$) . Indeed, grey-scale versions of moon paintings produce less constriction than their original versions ($t(2) = 2.96$, $p < 0.05$), and grey-scale versions of diffused light paintings produce less constriction than their original ($t(2) = 5.11$, $p < 0.001$) and inverted versions ($t(2) = 4.57$, $p < 0.01$). Therefore, in this condition, for all stimulus categories, the disruption of contextual cues alters pupillary response.

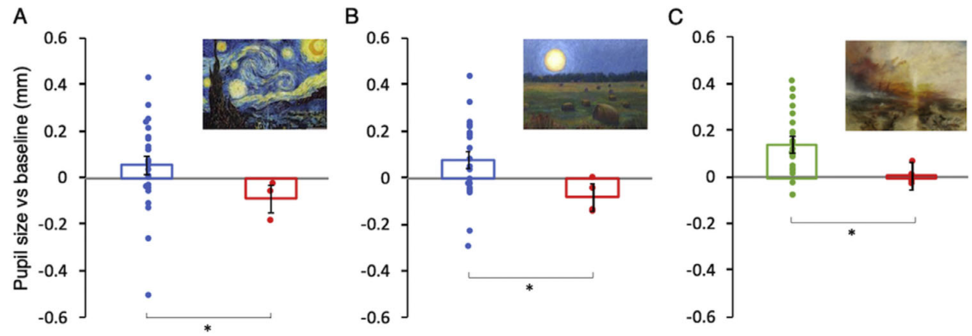

Figure 5. **Effects of subjective interpretation.** Single observer and average pupillary response μ_i (mm) for three stimuli subjected to three or more misinterpretations. Classification is based on the categorization of the light source made by the participants in the test. Blue: categorization as a moon, red: categorization as a sun, green: categorization as other. **(A)** Moon ($n = 25$): 0.06 ± 0.3; sun ($n = 3$): -0.09 ± 0.06. **(B)** Moon ($n = 24$): 0.8 ± 0.04); sun ($n = 4$): $-0.0.8 \pm 0.04$. **(C)** Other ($n = 24$): 0.14 ± 0.03); sun ($n = 4$): 0.01 ± 0.03. Error bars are the SE_i. Asterisks mark statistically significant comparisons between groups, non-parametric one-tailed Mann-Whitney test, $p < 0.05$*. Painting in **A** (image 17: *The Starry Night*, van Gogh, 1889) is in the public domain; copyright permission from the Author was obtained for painting in **B** (image 18: *Harvest moon*, Solberg, 2012); painting in **C** (image 23: *The slave ship* , Turner, 1840) is in the public domain.

Effects of subjective interpretation

Paintings are intrinsically complex stimuli, requiring a greater interpretative effort when compared with photographs and real-life scenes, leading to cases of ambiguous interpretation by observers. This is the reason for performing our main analysis based on individual observers' response to the categorization test (see Procedure). It is, however, interesting to look in more detail to the cases of ambiguous response. Figure 3A shows, image by image, not only the average response of conformant observations, but also displays the individual responses observed in the few cases of nonconforming categorizations. Inspection of Figure 3A clearly suggests that when observers classified a nominal sun painting as "other" (therefore, they did not see any light source) their pupil got a larger pupil size than that of those that had classified the same image as sun; whereas moon and diffused light paintings elicited a smaller pupil size in observers that had classified them as "sun" stimuli.

To test for the presence of the effect of subjective image interpretation, a nonparametric, one-tailed, Mann-Whitney ranking test was performed for data of all paintings that elicited differing responses in our experiment (in cases where only one misinterpretation occurred, the p value was directly determined as the ratio of the rank of the outlier and total number of subjects). Results show a significant effect for each case tested ($p < 0.05$). To assess the overall significance for the presence of an effect, individual p values were combined according to the Fisher's method (Mosteller & Fisher, 1948) yielding an overall p value < 0.0001. This is a strong indication for an influence of cognitive interpretation of a visual scene on the pupillary response of the observer.

Figure 5 shows, as an example, μ_i for the three most ambiguous stimuli of our set, each receiving 3 of 4 misclassifications in experiment 1, reported according to the categorization received ("sun," "moon," or "other").

Discussion

We show that artistic paintings, depicting scenes illuminated by light sources of different nature, such as the sun, moon, or containing a diffused lighting, although much less realistic than photographs in representing natural scenes and largely mediated by the artist's interpretation of reality and his technique, can differently modulate the pupillary response, according to the scene represented and not to their specific luminance or other low-level visual features.

In fact, despite that all paintings had the same mean luminance, when presented on a lighter background, paintings containing a light source produced less dilation than meaningless mean grey uniform-luminance rectangles, representing the control for dilation in this condition. In particular, paintings with the sun elicited a much smaller dilation than paintings with the moon, that in turn produced a lower dilation than paintings containing no visible light source.

This pattern of results does not depend on background luminance. When paintings are presented on a mean grey background, all produce constriction, although not expected from their average luminance that is equivalent to the background. This is in agreement with previous observations of the onset of changes in contrast, besides luminance, eliciting pupillary constriction (Naber et al., 2011, Naber & Nakayama, 2013). We find that the constriction induced by paintings containing a visible sun is larger than that produced by moon and diffused light paintings.

It is well known that the strength of pupillary response is larger for luminance changes occurring in the fovea (Clarke, Zhang, & Gamlin, 2003), and this raises the question of the role played by the higher values of luminance found in the vicinity of the fixation center in the case of sun and moon paintings. The fact that spatial distribution of luminance in the visual field and between image categories is not responsible for the observed differences between categories is demonstrated by three independent observations. First, when paintings are presented in the periphery, the same patterns of results are obtained: sun paintings produce less dilation than moon, diffused light and grey uniform control stimuli. This is in agreement with previous findings on photographic images (Binda et al., 2013b). Second, no correlation was found between pupil dilation and the local luminance measured at the center of suns or moons. Finally, the average luminance at the centers of sun and moon disks are compatible.

All the effects found for different stimulus categories do not depend on eye movements that have been shown to modulate pupil response (Gagl et al., 2011).

Our findings are in general agreement with those reported in the literature with non-painting stimuli (Binda et al, 2013b; Naber & Nakayama, 2013), but sun paintings produce a weaker effect compared with realistic pictures (Binda et al., 2013b). This might be the result of several factors, like differences in stimulus size and relative difference between luminance of stimuli and background. Our stimuli are also much more complex and may require higher cognitive load (Altschul et al., 2017; Tatler & Melcher, 2007), which is known to cause pupil dilation (Beatty, 1982; Hess & Polt, 1964; Just & Carpenter, 1993).

Results do not depend on the specific paintings chosen for the experiments, assigned to the three categories by the experimenters, and validated by all subjects in the categorization test. Although photograph categories chosen in similar studies comprise more or less homogeneous sets (see Binda et al., 2013b), here, paintings in the same category have been deliberately chosen to be as different as possible in style and period, to ensure the general validity of the findings. Despite this diversity, variability of responses to sun, moon, and diffused light paintings are the same, and, more importantly, they do not differ from the variability of responses to uniform grey control stimuli. This indicates that the pupil response is mainly driven by the scene depicted, overstepping differences in painting styles, artist's personal style, or his/her technique rendering of light sources.

Interesting results emerge also from the analysis of time variation of pupil size in experiments with light background. During the pre-stimulus interval there is a gradual increase of pupil diameter, possibly due to the effect of expectations (Irons, Jeon & Leber, 2017). During the first 500 ms after stimulus presentation, pupil diameter is mostly stable and equal for all the categories. This could be because the constriction that usually occurs when a stimulus appears (Naber & Nakayama, 2013; Naber et al., 2011; Privitera, Renninger, Carney, Klein, Aguilar, 2010) may be compensated by the dilation that should be produced by showing a stimulus darker than background. After this 500 ms period, pupil response starts to differ between categories. For all of them, though, there is a progressive increase of pupil size up until the end of the recording, consistent with the dilation effect due to cognitive load described in literature (Hess & Polt, 1964; Just & Carpenter, 1993; Kahneman & Beatty, 1966).

Interestingly, the variability of observers' responses to all categories of paintings also increases with time, being larger for sun paintings, whereas remaining more or less constant for the response to the uniform-grey control stimuli. Note that this same effect was also present in pupil responses to photographs (Binda et al., 2013b) or to words conveying a sense of brightness or darkness (Mathôt et al., 2017), although not analyzed or commented by the authors. We cannot be sure about the cause of this effect, but we could speculate that a number of different cognitive processes progressively set in while observers keep looking at the stimuli. This may include attention, recognition of elements in the painting, familiarity with the specific painting, aesthetic preference, memory, imagination, etc. All these factors, being different for each individual, produce a larger variability of responses than the one that could be generated by lower level perceptual visual mechanisms. This hypothesis is also in agreement with the observation that uniform-grey images, not involving such high-level processes, do not exhibit the same increase in variability.

Inverted paintings of the sun produce a larger pupil size than originals, despite sharing the same low-level features, such as luminance, contrast, chromatic contrast, and Fourier transform. This shows again that pupil amplitude is largely modulated by the observer's interpretation of the luminous objects rather than by its low-level features (Binda et al., 2013b; Naber & Nakayama, 2013). Image inversion is known to impair recognition performance of stimuli, such as pictures of faces, buildings, and cartoons (Naber & Nakayama, 2013; Scapinello & Yarmey, 1970; Strother et al., 2011; Valentine & Bruce, 1986; Van Belle, De Graef, Verfaillie, Rossion, & Lefevre, 2010; Yin, 1969). Therefore, by changing the complex relations between features of the paintings, the information about its content decreases, making it more difficult for the observer to use contextual cues to identify the source of light. A similar effect was found by Naber and Nakayama in computer generated images (Naber & Nakayama, 2013).

Grey-scale versions of sun paintings cause an even greater pupil size than originals, comparable to that produced by uniform-grey images, devoid of meaning, used as controls. Because chromatic content is a very important cue used for image interpretation (Goffaux et al., 2004, Greene & Oliva 2009; Oliva & Schyns, 2000; Oliva & Torralba, 2006; Steeves et al 2004), the fact that the absence of color in sun paintings increases pupil size is further proof of pupillary response being largely driven by interpretation of the light source. The suggestion that colored stimuli may produce different pupil response than their grey-scale versions was indeed previously advanced, although not systematically investigated (Snowden et al., 2016).

The grey-scale versions of our sun stimuli also cause a larger pupil size than inverted versions, suggesting that color cues are even more important than spatial organization for the identification of the light source.

Note that the presentation of each painting in three different versions does not affect pupil responses, as expected with multiple exposures to the same stimulus (Yoshimoto et al., 2014), probably because the three versions are not perceived as repetitions of the same stimulus.

The chromatic structure of artistic compositions mostly follows the statistical features of the natural environment (Montagner et al., 2016). Therefore, blue colors are generally used in night scenes representations, whereas yellow-reddish chromaticities are used in rendering daylight scenes. Thus, different response to moon and sun paintings might be ascribed to their different chromatic contents. However, the results of this work imply that the presence of an object interpretable as a light source plays a crucial role in scene reconstruction. Indeed, diffused-light paintings endowed with the same yellow-reddish chromaticities of sun paintings, but no visible light source, produce distinguishably larger pupil size.

Perhaps the most convincing evidence presented in this work for the crucial role of image interpretation in pupillary response, is the strong relationship observed between pupil diameter of observers and their subjective interpretation of the light source. The same painting is capable of eliciting constriction in observers who see it as a sun representation and dilation in those who see it as a moon.

Other authors have tried to explain why showing images with a sun produces more constriction than images of the same luminance with different lighting structure, and we can reasonably presume that these explanations may hold also for the effects found with our paintings. A potential explanation is that the subjective perception of increased brightness reduces pupil size, as found with illusions by Laeng and Endestad (2012) and Suzuki et al., (2019) with psychophysical methods. Nevertheless, Binda et al., (2013b), by using a rating method of their stimuli, did not find this correlation. Moreover, Naber and Nakayama (2013) demonstrated that even cartoon depictions of the sun, appearing no brighter than cartoon depictions of the moon, can result in pupil constrictions. Another proposed explanation is based on different spatial distribution of attention across image categories, as it is known that attention strongly affects pupil size (Binda et al., 2013a). The observer's attention might be focusing more on the brighter regions of the sun pictures and spread more evenly in other images. However, this hypothesis has been ruled out by Binda et al., (2013b), showing that photographs of the sun cause constriction even when the observer's attention is directed to performing a different task. An explanation that still remains open after the present work is that of a protective behavior against a potentially harmful light level triggered by high-level interpretation of a very luminous object (Binda et al., 2013b; Laeng & Endestad, 2012; Naber & Nakayama, 2013; Suzuki et al., 2019). In other words, we can hypothesize that our system initiates a defense response to the powerful light induced by the sun, even if it is just depicted in a painting.

All evidences presented in this work converge with the results of previous studies in suggesting a top-down control on the pupillary light reflex (Becket Ebitz & Moore, 2019; Binda & Murray, 2015a). The neural pathways underling this high-level modulation of PLR cannot be identified with certainty, but some potentially relevant circuits have already been identified. It is well established that pupillary constriction results from the activation of the subcortical Edinger-Westphal nucleus (EW) (Gamlin & Clarke, 1995), and there are some known modulatory inputs from cortical areas to this circuit. First, EW activity is enhanced by inputs from the visual cortex (Becket Ebitz & Moore, 2017; Binda & Gamlin, 2017) and the superior colliculus (Gamlin, 2006; Joshi & Gold, 2019; Joshi, Li, Kalwani, & Gold, 2016; Wang & Munoz, 2015; Wang & Munoz, 2012). Other possible inputs to the PLR could come directly from the prefrontal cortex, in particular from the frontal eye field (FEF), or indirectly through the extrastriate cortex, the oculomotor regions in the parietal cortex and the superior colliculus that are modulated by FEF (Becket Ebitz & Moore, 2017). EW nucleus also receives inhibitory input from the sympathetic system through projections from locus coeruleus (Joshi et al., 2016; Peinkhofer et al., 2019) and the hypothalamus that are potentially under cortical control (Aston-Jones & Cohen, 2005). A reduction of this inhibitory inputs could result in a pupillary constriction (Joshi & Gold, 2019; Wilhelm, 2002).

Conclusions

The present work provides further evidence for the influence of high-level visual processing on the modulation of pupil response, corroborating an existing body of evidence. However, our specific choice of paintings as stimuli allows to push the exploration of the involved top-down mechanisms toward the even higher-level cognitive processing involved in aesthetic experience, imagination and memory. This reaches a point where the very same image can produce opposite responses depending on the individual subjective interpretation and visual awareness.

Overall, this suggests that variations of pupil diameter can be an effective probe into cortical processing, making pupillometry a useful tool for the study of high-level vision and cognition.

Keywords: pupillometry, pupillary response modulation, high-level visual processing, artistic representation of light, aesthetic experience

Acknowledgments

The authors thank Paola Binda e Marco Turi for helpful discussion and comments on the data. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No 832813 GenPercept "Spatio-temporal mechanisms of generative perception").

Commercial relationships: none.
Corresponding author: Maria Michela Del Viva.
Email: maria.delviva@unifi.it.
Address: Department of Neurofarba, University of Florence, Florence, Italy.

References

Altschul, D., Jensen, G., & Terrace, H. (2017). Perceptual category learning of photographic and painterly stimuli in rhesus macaques (Macaca mulatta) and humans. *PLoS One, 12*(9), e0185576.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403–450.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276–292.

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5*, 371–372.

Becket Ebitz, R., & Moore, T. (2017). Selective modulation of the pupil light reflex by microstimulation of prefrontal cortex. *Journal of Neuroscience, 37*(19), 5008–5018.

Becket Ebitz, R., & Moore, T. (2019). Both a gauge and a filter: Cognitive modulations of pupil size. *Frontiers in Neurology, 9*, 1190.

Bierstadt, A., Painter. 1850. *Aurora* (Painting). March 28, 2019. Retrieved from: https://wikioo.org/it/paintings.php?refarticle=8BWL6X&titlepainting=Sunrise&artistname=Albert%20Bierstadt.

Binda, P., & Gamlin, P. D. (2017). Renewed attention on the pupil light reflex. *Trends in Neurosciences, 40*(8), 455–457.

Binda, P., & Murray, S. O. (2015a). Spatial attention increases the pupillary response to light changes. *Journal of Vision, 15*(2), 1.

Binda, P., & Murray, S. O. (2015b). Keeping a large-pupilled eye on high-level visual processing. *Trends in Cognitive Sciences, 19*(1), 1–3.

Binda, P., Pereverzeva, M., & Murray, S. O. (2013a). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience, 33*(5), 2199–2204.

Binda, P., Pereverzeva, M., & Murray, S. O. (2013b). Pupil constrictions to photographs of the sun. *Journal of Vision, 13*(6), 8.

Binda, P., Pereverzeva, M., & Murray, S. O. (2014). Pupil size reflects the focus of feature-based attention. *Journal of Neurophysiology, 112*(12), 3046–3052.

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.

Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. *Trends in Cognitive Sciences, 18*(7), 370–375.

Clarke, R. J., Zhang, H., & Gamlin, P. D. R. (2003). Characteristics of the pupillary light reflex in the alert rhesus monkey. *Journal of Neurophysiology, 89*(6), 3179–3189.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Science* (2nd Edition). In *Statistical Power Anaylsis for the Behavioral Sciences*. Milton Park, Abingdon, UK: Lawrence Erlbaum Associates.

Creti, D., Painter. 1711. *Astronomical Observations: the Moon (Painting)*. Pinacoteca Vaticana: Rome, Italy. March 28, 2019. Retrieved from: https://commons.wikimedia.org/wiki/File:Donato_Creti_-_Astronomical_Observations_-_01_-_Sun.jpg.

De Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *Proceedings of the National Academy of Sciences of the United States of America, 111*(5), E618–E625.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences of the United States of America, 105*(5), 1704–1709.

Fahle, M. W., Stemmler, T., & Spang, K. M. (2011). How much of the "unconscious" is just pre-threshold? *Frontiers in Human Neuroscience, 5*, 120.

Gagl, B., Hawelka, S., & Hutzler, F. (2011). Systematic influence of gaze position on pupil size measurement: analysis and correction. *Behavior Research Methods, 43*(4), 1171–1181.

Gamlin, P. D. R. (2006). The pretectum: Connections and oculomotor-related roles. *Progress in Brain Research, 151*, 379–405.

Gamlin, P. D., & Clarke, R. J. (1995). The pupillary light reflex pathway of the primate. *Journal of the American Optometric Association, 66*(7), 415–418.

Gercken, G., Painter. 2012. *Rural sunrise* (Painting). March 28, 2019. Retrieved from: https://fineartamerica.com/featured/rural-sunrise-graham-gercken.html.

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P. G., & Rossion, B. (2004). Diagnostic colors contribute to the early stages of scene categorization: Behavioral and neurophysiological evidence. *Journal of Vision, 12*(6), 878–892.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science, 21*(2), 90–95.

Graham, D. J., & Field, D. J. (2008). Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision, 21*(1–2), 149–164.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding: Research Article. *Psychological Science, 20*(4), 464–472.

Henderson, R. R., Bradley, M. M., & Lang, P. J. (2014). Modulation of the initial light reflex during affective picture viewing. *Psychophysiology, 51*(9), 815–818.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science, 132*(3423), 349–350.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143*(3611), 1190–1192.

Irons, J. L., Jeon, M., & Leber, A. B. (2017). Pre-stimulus pupil dilation and the preparatory control of attention. *PLoS One, 12*(12), e0188787.

Joshi, S., & Gold, J. I. (2019). Pupil size as a window on neural substrates of cognition. *Trends in Cognitive Sciences, 24*(6), 466–480.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron, 89*(1), 221–234.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Expérimentale, 47*(2), 310–339.

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly Journal of Experimental Psychology, 64*(10), 1971–1989.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*(3756), 1583–1585.

Kimura, E., Abe, S., & Goryo, K. (2014). Attenuation of the pupillary response to luminance and color changes during interocular suppression. *Journal of Vision, 14*(5), 14.

Kohn, M., & Clynes, M. (1969). Color dynamics of the pupil. *Annals of the New York Academy of Sciences, 156*(2), 931–950.

Kuchinke, L., Trapp, S., Jacobs, A. M., & Leder, H. (2009). Pupillary responses in art appreciation: effects of aesthetic emotions. *Psychology of Aesthetics, Creativity, and the Arts, 3*(3), 156–163.

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

Laeng, B., & Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological Science, 25*(1), 188–197.

Loewenfeld, I. (1993). *The pupil: Anatomy, physiology, and clinical applications*. Detroit, MI: Wayne State University Press.

Marković, S. (2010). Aesthetic experience and the emotional content of paintings. *Psihologija, 43*(1), 47–64.

Marković, S. (2011). Perceptual, semantic and affective dimensions of experience of abstract and representational paintings. *Psihologija, 10*(7), 1230.

Marković, S. (2012). Components of aesthetic experience: Aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *I-Perception, 3*(1), 1–17.

Marković, S., & Radonjić, A. (2008). Implicit and explicit features of paintings. *Spatial Vision, 21*(3–5), 229–259.

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, 50*(1), 94–106.

Mathôt, S., Grainger, J., & Strijkers, K. (2017). Pupillary responses to words that convey a sense of brightness or darkness. *Psychological Science, 18*(8), 1116–1124.

Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2013). The pupillary light response reveals the focus of covert visual attention. *PLoS One, 8*(10), e78168.

Monet, C., Painter. 1888. *Landscape at Giverny* (Painting). March 28, 2019. Retrieved from: https://commons.wikimedia.org/wiki/File:Monet_-_landscape-at-giverny.jpg.

Montagner, C., Linhares, J. M. M., Vilarigues, M., & Nascimento, S. M. C. (2016). Statistics of colors in paintings and natural scenes. *Journal of the Optical Society of America A Optics, Image Science, and Vision, 33*(3), A170–A177.

Mosteller, F., & Fisher, R. (1948). Questions and Answers. *The American Statistician, 2*(5), 30–31.

Naber, M., & Nakayama, K. (2013). Pupil responses to high-level image content. *Journal of Vision, 13*(6), 7.

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology, 4*, 919.

Naber, M., Frässle, S., & Einhäuser, W. (2011). Perceptual rivalry: Reflexes reveal the gradual nature of visual awareness. *PLoS One, 6*(6), e20910.

Nascimento, S. M. C., Linhares, J. M. M., Montagner, C., João, C. A. R., Amano, K., Alfaro, C., . . . Bailão, A. (2017). The colors of paintings and viewers' preferences. *Vision Research, 130*, 76–84.

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41*(2), 176–210.

Oliva, A., & Torralba, A. (2006). Chapter 2 Building the gist of a scene: the role of global image features in recognition. *Progress in Brain Research, 155*, 23–36.

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology, 83*(1), 56–64.

Peinkhofer, C., Knudsen, G. M., Moretti, R., & Kondziella, D. (2019). Cortical modulation of pupillary function: Systematic review. *PeerJ Computer Science, 7*, e6882.

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision, 10*(10), 3.

Sahraie, A., & Barbur, J. L. (1997). Pupil response triggered by the onset of coherent motion. *Graefe's Archive for Clinical and Experimental Ophthalmology, 235*(8), 494–500.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8*(2), 597–599.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21*(6), 329–331.

Snowden, R. J., O'Farrell, K. R., Burley, D., Erichsen, J. T., Newton, N. V., & Gray, N. S. (2016). The pupil's response to affective pictures: Role of image duration, habituation, and viewing mode. *Psychophysiology, 53*(8), 1217–1223.

Solberg, B., Painter. 2012. *Harvest moon* (Painting). March 28, 2019. Retrieved from: http://www.bsolbergart.com/gallery/IMG79.htm?reload.

Sperandio, I., Bond, N., & Binda, P. (2018). Pupil size as a gateway into conscious interpretation of brightness. *Frontiers in Neurology, 9*, 1070.

Steeves, J. K. E., Humphrey, G. K., Culham, J. C., Menon, R. S., Milner, A. D., & Goodale, M. A. (2004). Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *Journal of Cognitive Neuroscience, 16*(6), 955–965.

Strother, L., Mathuranath, P. S., Aldcroft, A., Lavell, C., Goodale, M. A., & Vilis, T. (2011). Face inversion reduces the persistence of global form and its neural correlates. *PLoS One, 6*(4), e18705.

Suzuki, Y., Minami, T., Laeng, B., & Nakauchi, S. (2019). Colorful glares: Effects of colors on brightness illusions measured with pupillometry. *Acta Psychologica (Amsterdam), 198*, 102882.

Tatler, B. W., & Melcher, D. (2007). Pictures in mind: Initial encoding of object properties varies with the realism of the scene stimulus. *Perception, 36*(12), 1715–1729.

Tkacz-Domb, S., & Yeshurun, Y. (2018). The size of the attentional window when measured by the pupillary response to light. *Scientific Reports, 8*(1), 11878.

Turner, W., Painter. 1840. The Slave Ship (Painting). Museum of Fine Arts: Boston. July 28, 2020. Retrieved from: https://commons.wikimedia.org/wiki/File:Slave-ship.jpg.

Ukai, K. (1985). Spatial pattern as a stimulus to the pupillary system. *Journal of the Optical Society of America A, 2*(7), 1094–1100.

Unsworth, N., Robison, M. K., & Miller, A. L. (2018). Pupillary correlates of fluctuations in sustained attention. *Journal of Cognitive Neuroscience, 30*(9), 1241–1253.

Valentine, T., & Bruce, V. (1986). The effect of race, inversion and encoding activity upon face recognition. *Acta Psychologica (Amsterdam), 61*(3), 259–273.

Van Belle, G., de Graef, P., Verfaillie, K., Rossion, B., & Lefèvre, P. (2010). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision, 10*(5), 10.

Van Gogh, V.W., Painter. 1889. *The Starry Night (Painting)*. New York City, NY: Museum of Modern Art (MoMA). March 28, 2019. Retrieved from: https://commons.wikimedia.org/wiki/File:VanGogh-starry_night.jpg#filelinks.

Wang, C. A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology, 33*, 134–140.

Wang, C. A., Boehnke, S. E., White, B. J., & Munoz, D. P. (2012). Microstimulation of the monkey superior colliculus induces pupil dilation without evoking saccades. *Journal of Neuroscience, 32*(11), 3629–3636.

Wilhelm, B. J. (2002). Pupil response components: studies in patients with Parinaud's syndrome. *Brain, 125*(Pt 20), 2296–2307.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Yoshimoto, S., Imai, H., Kashino, M., & Takeuchi, T. (2014). Pupil response and the subliminal mere exposure effect. *PLoS One, 9*(2), e90670.

Zavagno, D., Tommasi, L., & Laeng, B. (2017). The eye pupil's response to static and dynamic illusions of luminosity and darkness. *I-Perception, 8*(4), 2041669517717754.