

## RESEARCH ARTICLE

## Spectral dependency of the human pupillary light reflex. Influences of pre-adaptation and chronotype

Johannes Zauner<sup>1\*</sup>, Herbert Plischke<sup>1‡</sup>, Hans Strasburger<sup>2‡</sup>**1** Munich University of Applied Sciences, Munich, Germany, **2** Institute of Medical Psychology, Ludwig-Maximilians-Universität, Munich, Germany

‡ HP and HS are Joint Senior Supervisors.

\* [zauner@me.com](mailto:zauner@me.com)

## OPEN ACCESS

**Citation:** Zauner J, Plischke H, Strasburger H (2022) Spectral dependency of the human pupillary light reflex. Influences of pre-adaptation and chronotype. *PLoS ONE* 17(1): e0253030. <https://doi.org/10.1371/journal.pone.0253030>

**Editor:** Manuel Spitschan, University of Oxford, UNITED KINGDOM

**Received:** May 22, 2021

**Accepted:** October 30, 2021

**Published:** January 12, 2022

**Copyright:** © 2022 Zauner et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data are available from DOI: [10.17605/OSF.IO/AR2UD](https://doi.org/10.17605/OSF.IO/AR2UD).

**Funding:** This work was financially supported by the Hochschule München University of Applied Sciences (MUAS) and the German Research Foundation (DFG) through the "Open Access Publishing" program. MUAS and DFG had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Non-visual photoreceptors (ipRGCs) and rods both exert a strong influence on the human pupil, yet pupil models regularly use cone-derived sensitivity as their basis. This inconsistency is further exacerbated by the fact that circadian effects can modulate the wavelength sensitivity. We assessed the pupillary reaction to narrowband light stimuli in the mesopic range. Pupil size for eighty-three healthy participants with normal color vision was measured in nine experimental protocols with varying series of continuous or discontinuous light stimuli under Ganzfeld conditions, presented after 90 seconds of dark adaptation. One hundred and fifty series of stimulation were conducted across three experiments, and were analyzed for wavelength-dependency on the normalized pupillary constriction (nPC), conditional on experimental settings and individual traits. Traits were surveyed by questionnaire; color vision was tested by *Ishihara plates* or the *Lanthony D15* test. Data were analyzed with generalized additive mixed models (GAMM). The normalized pupillary constriction response is consistent with L+M-cone derived sensitivity when the series of light stimuli is continuous, i.e., is not interrupted by periods of darkness, but not otherwise. The results also show that a mesopic illuminance weighing led to an overall best prediction of pupillary constriction compared to other types of illuminance measures. IpRGC influence on nPC is not readily apparent from the results. When we explored the interaction of chronotype and time of day on the wavelength dependency, differences consistent with ipRGC influence became apparent. The models indicate that subjects of differing chronotype show a heightened or lowered sensitivity to short wavelengths, depending on their time of preference. IpRGC influence is also seen in the post-illumination pupil reflex if the prior light-stimulus duration is one second. However, shorter wavelengths than expected become more important if the light-stimulus duration is fifteen or thirty seconds. The influence of sex on nPC was present, but showed no interaction with wavelength. Our results help to define the conditions, under which the different wavelength sensitivities in the literature hold up for narrowband light settings. The chronotype effect might signify a mechanism for strengthening the individual's chronotype. It could also be the result of the participant's prior exposure to light (*light history*). Our explorative findings for this effect demand replication in a controlled study.

## Introduction

The human pupil is of interest to various research fields, such as vision research, neurobiology, ophthalmology, and psychology. Pupil constriction and size are controlled by parasympathetic pathways, and pupil dilation through sympathetic pathways [1, 2]. Pupillary reaction thus acts as a window into the autonomous nervous system. Diagnostic methods were developed around various aspects of the pupil's behavior and a vast body of research is dedicated to the pupillary reaction to light. One of the most fundamental aspects in that context is the dependency of the pupillary light reflex (PLR) on the spectral composition of light.

Up until now, the literature is somewhat divided on what the appropriate spectral weighing function should be for capturing the effects of light on pupil size and constriction. Pupil models derived from research using white light stimuli, i.e., polychromatic light spectra, generally employ stimulus luminance or retinal illuminance as predictors. Watson and Yellott [3] in 2012 reviewed eight pupil models published between 1926 and 1999. They created a ninth model by incorporating elements of the previous ones. What is important in our case is that all these models, including the newly created one, work with standard photometric dimensions, and spectral calculations are thus based on the  $V(\lambda)$  weighing function, regardless of stimulus characteristics. However, the  $V(\lambda)$  function may not be the best-suited for predicting pupil size in all cases.

The  $V(\lambda)$  function peaks at around 555 nm wavelength. Yet studies using narrowband light stimuli generally report a maximum sensitivity for PLR at around 480–510 nm. Early publications to this effect are notably by Wagman and Gullberg [4] in 1942, by Alpern and Campbell [5] in 1962, and the often-cited paper by Bouma [6] in 1962. These authors attributed the PLR's spectral dependency mainly to the characteristics of rod photoreceptors, but possibly also to those of short-wavelength cone receptors. Adrian [7] in 2003 tried to explain the apparent blue shift in sensitivity, reporting that, with appropriate adjustments, stimulus luminance was sufficient for explaining the PLR. According to the paper, a  $V_{10}(\lambda)$  spectral weighing function which is based on the larger 10-degree field should be used instead of the usual 2-degree field (on which  $V(\lambda)$  is based) which, according to the author, is valid for the often-used Ganzfeld conditions. Furthermore, Adrian argues, mesopic conditions should be adequately addressed by using  $V_{eq}(\lambda)$  weighting functions as intermediaries between the photopic  $V_{10}(\lambda)$  and the scotopic  $V'(\lambda)$  functions. With these adjustments, pupil size depended linearly on (adjusted) luminance. While Adrian [7] performed his analysis on data from a single subject and, very likely, did not know about the spectral-temporal changes of the PLR as found in later studies, his argument about the appropriate luminance function remains valid, even when it was not widely adopted.

With the discovery of intrinsically photosensitive retinal ganglion cells (ipRGCs) at the beginning of the 21st century [8], and their potential to drive pupillary constriction [9], researchers began looking for their influence in PLR experiments. In 2007, Gamlin et al. [10] made the ipRGC influence on pupilloconstriction evident in the macaque with pharmacological blockade of rod and cone photoreceptors. They also showed ipRGC dependency for the post-illumination pupil reflex (PIPR) of humans. Zaidi et al. [11], also in 2007, confirmed the human pupilloconstriction's dependency on ipRGC sensitivity in a blind person who lacked an outer retina. In 2009, Mure et al. [12] found ipRGC-dependent pupilloconstriction in the PIPR of sighted humans and, more importantly, during more prolonged light stimulation than used in previous studies (five minutes, compared to a few seconds in earlier studies). In 2010, McDougal and Gamlin [13] investigated the contributions of individual receptors to the PLR, for stimulus durations between 1 and 100 seconds. They did this through parameter fitting in a custom-derived function based on theoretical assumptions. They found that rods play an

essential role for most stimulus durations, while cones contribute only minimally after just 10 seconds of stimulus onset. At about 18 seconds after stimulus onset, the ipRGC sensitivity curve already provided the most prominent contribution. It has not been confirmed whether the stated time intervals for the receptor types are valid in humans. In 2012, Gooley et al. [14] used stimulus durations from 2 to 90 minutes, showing for the PLR, among others, the dominance of shorter wavelengths at around 490 nm, compared to the brighter settings around 555 nm, especially for longer stimulus durations.

Spitschan [15] showed the inadequacy of using  $V(\lambda)$ -derived luminance instead of a melanopic weighing function by reanalyzing the data from Bouma [6]. He estimated the possible error for the stimulus calculation to predict pupil size at about one order of magnitude for typical white-light sources, with the actual error and relevance depending on the application's circumstances. A recent paper by Zandi et al. [16] looked at the prediction accuracy of the models of Watson and Yellott [3]. They found that the prediction error was greatest for chromatic spectra, especially for longer exposure times. In the four cases of short chromatic light stimuli of equal luminance (1 second, 100 cd/m<sup>2</sup>), Zandi et al. [16] found that the significant differences between wavelengths were relatively small. For more extended stimulation periods of 60 and 300 seconds, a luminance-based approach led to considerable error. Together with findings from polychromatic spectra, the authors therefore advocated for the inclusion of time and spectrum in pupil models. Rao et al. [17] suggested a pupil model based on a combined measure of photopic luminance and an action spectrum based on melatonin suppression ( $C(\lambda)$ , *circadian sensitivity*) that takes spectral differences of the stimulus into account. While Spitschan [15] calls it a good start, both he and Zandi et al. [16] find shortcomings, ranging from the narrow luminance range under which the model parameters were estimated [15], the use of the non-standardized circadian sensitivity function, the relatively small expected improvement in prediction accuracy, to the lack of (stimulus) time as part of the model [16].

In summary, by addressing the effects of time since light onset, the ipRGC's influence on PLR has become evident for narrowband stimuli [18]. It is also accepted that the many metrics of the PLR, like latency to constriction, peak constriction, rhodopsin- and melanopsin-mediated PIPR amplitude, or melanopsin-mediated PIPR amplitude, differ in their retinal circuitry and thus their spectral sensitivity [15, 19]. These findings sparked the development of diagnostic methods [20] for assessing the health of the nonvisual pathway [21] and circadian system [22], and spotting specific pathologies early on, such as glaucoma [23] and diabetic retinopathy [24]. However, what is still lacking is knowledge about why the PLR in terms of pupil size and constriction during light stimulation seems to differ in spectral sensitivity between narrowband light and polychromatic (white) light. The applicability of prior findings based on narrowband stimuli may be limited for polychromatic or even narrowband settings. Furthermore, there is still little research into differences of spectral dependency beyond comparing particular wavelengths, mostly those of the blue and red peaks. Most of the studies that use a broad range of wavelengths are based on a comparatively small sample size. The largest of the studies reviewed above had twenty-four participants [14], while the median is only four participants [4–7, 10–14]. While this is enough for investigating the fundamentals of the PLR's wavelength sensitivity, exploring distributional relationships will require more participants.

We initially set out to take a comparatively large sample of participants for looking at the conditional spectral dependency of pupilloconstriction. Since the ipRGC influence on PLR is mainly present for long-lasting narrowband light stimuli, we believed that, basically, pupilloconstriction would approximately follow the ipRGC sensitivity curve for a series of continuously applied narrowband light stimuli that change in wavelength over several minutes. Preliminary measurements supported this view and further indicated a dependence on sex that we included as an exploratory variable in the main experiment since there is also some

support for sex differences in the literature [25]. We were further interested in the influence of chronotype and time of day (cf. [22, 26]); findings from Zele et al. [27] suggested that there is a circadian variation in the wavelength-dependency, as demonstrated by pupil reactions to red and blue stimuli. In our main experiment with a balanced design, wavelengths around the V ( $\lambda$ ) function's peak led to the highest constriction. Based on these findings, we designed an exploratory, second experiment, to test whether our findings held up with a changed stimulus series. The findings were indeed replicated, but only in the case when light remained on continuously. When light was discontinuous, i.e., when short intervals of darkness were introduced between the light steps, shorter wavelengths became more influential, in line with published literature. The newly introduced periods of darkness in the second experiment additionally allowed spectral analysis of the post-illumination pupil reflex (PIPR). In a third experiment, we therefore investigated (1) the spectral dependency of the PLR for longer-lasting light steps with short, intermediate periods of darkness, as well as (2) the PIPR for short light steps with more prolonged intermediate periods of darkness.

We believe our findings will help bridge the gap between the canonical divide of luminance-based versus short-wavelength-based pupil models by showing that the PLR's response to narrowband light stimuli can go both ways, depending on pre-adaptation to light. We also show the influence of chronotype on pupilloconstriction, depending on the time of day. Statistical analyses employ generalized additive models (GAMs), which to our knowledge have not been used in pupillometry before. They appear particularly well suited for disentangling the complex coaction of influences.

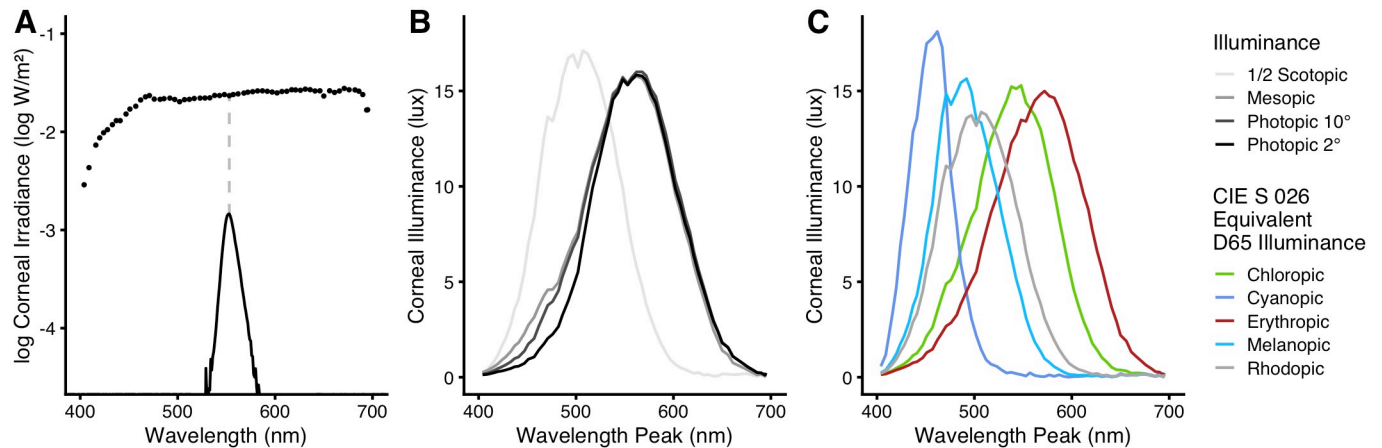
## Materials and methods

### Participants

We recruited a total of 83 young, healthy participants across three experiments through bulletin boards and announcements at the Munich University of Applied Sciences (45 females and 38 males; age: median 26 yr., range 18–36; chronotype score:  $50 \pm 10.6$  on the Morningness-Eveningness Questionnaire, where a score between 40 to 60 signifies an intermediary chronotype, below 40 an evening type, and above 60 a morning type [28]). All procedures were approved by the Ethics Committee of the Munich University of Applied Sciences. Exclusion criteria were age (above 39), issues of psychiatric or neuronal health, addiction diseases, regular intake of stimulants or sedatives, acute jetlag or shiftwork (during the past three months), ocular diseases, or epilepsy. Participants were instructed to refrain from caffeine intake prior to the experimental sessions. We further checked participants for normal color vision (Lanthony D15 test [29] or Ishihara plates [30]). Additional participants were tested but their results not included in the analysis [31]; these participants either only took part in the preliminary measurements (see below), did not fit the inclusion criteria, or were tested after the experiment's cut-off date.

### Apparatus

The light stimuli were presented in a Ganzfeld dome setup (Light Dome Model XE 509, *Monocrom*, Stockholm). The light source was a 100-Watt Xenon lamp. Full-spectrum light (6000 Kelvin correlated color temperature) was filtered through a stepper-controlled monochromator and led into the dome's interior by a diffusor which was positioned above and behind participants' eyes and thus not visible during the experiment. Light onset and offset were controlled by a shutter mechanism, located between the monochromator and the diffusor. Shutter response time, i.e., the time between open and closed states, was about 100 ms. Changes in peak wavelength occurred at a speed of about 35 nm/s. An *Arduino MEGA-2560*



**Fig 1. Spectral irradiance measurements and calculated illuminance values at the eye level.** All displayed values are based on spectral irradiance measurements with a field-of-view restriction according to the CIE S 026 standard [32]. In our case, these measurements are 24% lower than those of the unobstructed sensor diffusor. Differences in irradiance measurements from the mean were smaller than 5%, or about 1% on the log irradiance scale. Differences in wavelength peak were between 0 to 1 nm. (A) Dots show the log of the corneal irradiance (in log W/m<sup>2</sup>) of exemplary steps at their respective peak wavelength. Each dot represents a narrowband peak, similar to the spectral example distribution with a peak at 553 nm (the peak is indicated by the gray dashed line). (B) Corneal illuminance values for the narrowband light steps at each peak wavelength. Besides the *standard* photopic CIE-1931 V(λ) weighing for the 2° observer (black line), illuminance values were calculated based on the weightings of V<sub>10</sub>(λ) (10° observer, photopic vision, dark grey), of V'(λ) (scotopic vision, light grey), and of V<sub>eq</sub>(λ) (mesopic vision, medium grey). To show the scotopic illuminance values along with the other curves, they are drawn at half their original value (light-gray curve). (C) Corneal alpha-opic equivalent-daylight illuminance for all receptor types according to the CIE S 026 standard [32].

<https://doi.org/10.1371/journal.pone.0253030.g001>

embedded computer with custom-written software controlled the monochromator and shutter.

The setup allows generating narrowband light, with spectral peaks at between 400 nm and 700 nm, and a width (full width at half maximum) of 12 nm ± 1 nm (Fig 1). We used 61 or 31 peaks, depending on the experiment. The radiant output from the light source is dependent on wavelength (a limitation of the apparatus) and irradiance increases by about one order of magnitude from 404 nm to 450 nm; above that and up to 690 nm, irradiance is comparatively constant at a level of  $-1.62 \pm 0.04 \log_{10}(\text{W}/\text{m}^2)$ . To accommodate for that dependency, in particular the lower irradiance levels for the wavelengths at and below 450 nm, we included the  $\log_{10}$  of the irradiance as a covariate in all statistical analyses when it was not already included implicitly, i.e., as following from illuminance. Measurements of spectral irradiance were taken at the eye level under the light dome. Measurements further included a human field-of-view restriction suggested by CIE Standard S 026/E:2018 [32]. Therefore, the irradiance measurements represent the corneal level, as do the various illuminance calculations based on those measurements (Fig 1). Unobstructed measurements, i.e., *regular* irradiance measurements, were about 24% higher. Plots of all spectral peaks are available in the S13 Fig. We measured the spectral irradiance of each wavelength step multiple times before each experiment to ensure stability in irradiance and wavelength peak. Differences in irradiance were relatively consistent across the spectrum and were smaller than 5%, which equals about 1% on the log irradiance scale from the values shown in Fig 1A (see S14 Fig for a regular scale). Differences in wavelength peak were between 0 to 1 nm, which is within the measurement accuracy of the spectrometer. The light source was not monitored during the experiments other than to ensure that it performed on a constant operating temperature.

Park et al. [21] summarized stimuli sizes from three studies [10, 13, 33], stating that melanopsin responses were “observed in a range between ~11 to >14 log quanta/cm<sup>2</sup>/s for a



470-nm light". Our stimulus for 471 nm lies at 12.75 quanta/cm<sup>2</sup>/s at the corneal level. However, this is not directly comparable to the summary from Park et al. [21] since those studies used focused light projections. Park et al. [21] converted these measures to luminance values (-0.7 to 2.3 log cd/m<sup>2</sup>). In our setup, the lower limit of that study translates to -2.3 log W/m<sup>2</sup>, compared to our corneal irradiance of -1.6 log W/m<sup>2</sup> at 471 nm.

## Measurement equipment

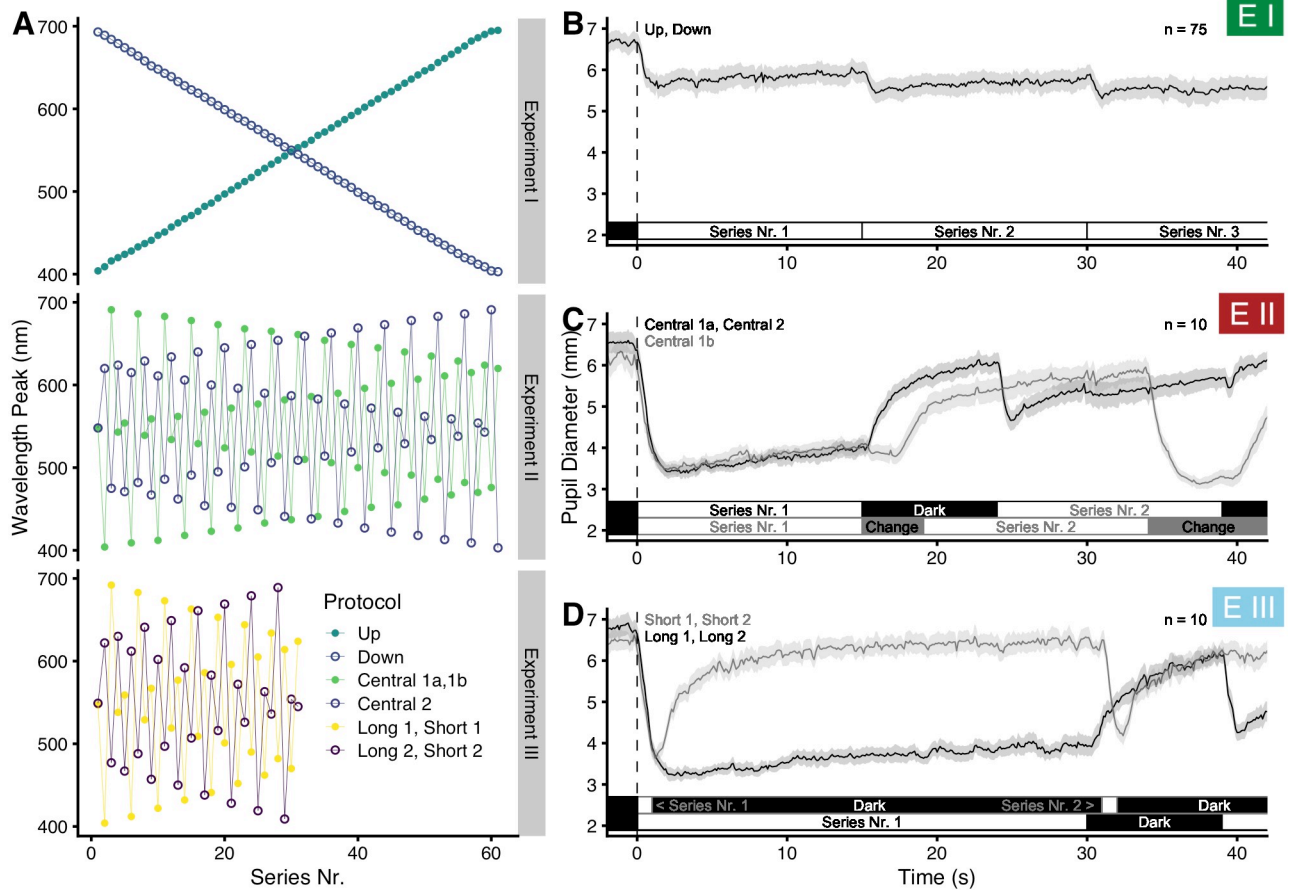
An eye tracker (*Dikablis Professional Glasses, Ergoneers GmbH, Munich*) recorded the pupils by a dual infrared camera setup for simultaneous measurement of the two eyes, at a sampling rate of 60 Hz. Cameras in that setup are located at the end of flexible arms and are adjustable for an optimal view of the pupil and distance from eye level. Each camera has a resolution of 384 by 288 pixel. Camera pictures were analyzed in real time while recording through the *D-Lab 3.5* software (*Ergoneers GmbH*) on a connected personal computer. The software extracted several pupil parameters for each eye in pixel units, including pupil area, height, and width. These time-stamped variables were then exported to a comma-separated-value (CSV) file for later analysis. Spectral irradiance measurements with a 1-nm resolution were performed using a *JETI Specbos 1201* spectroradiometer (*JETI Technische Instrumente GmbH, Jena*; see Fig 1), with the *JETI LiVal V6.14.2* software running on a connected personal computer. The spectroradiometer's relative measurement accuracy is 2%.

## Experimental design

**Preliminary measurements.** In the relevant literature, the employed periods of dark adaptation (DAP) before light onset vary widely, if reported at all. They range from no-DAP [10] to 40 minutes [12]; the use of an in-between value of 20 minutes for full dark adaptation is recommended by Kelbsch et al. [19]. This value is commonly used when there is a main stimulus following the DAP. Since we planned on using comparatively long experimental light durations of around 15 minutes containing a chain of wavelength peaks, all equally important, we did not want dark adaptation to overly influence the spectral dependency during the first light steps of the experiment, compared to the last steps, where light adaptation would have happened for several minutes, regardless of the dark adaptation period before the first presented stimulus. We tried periods of 90 seconds, 180 seconds, and 900 seconds on two participants with the *Up* protocol in three experimental runs, which is described under *Experiment I* below. The longer the DAP, the greater was pupilloconstriction during the first third of the steps (see S1 Fig). We compared the results to a second run of the *Up* protocol directly after the first one. In this second run, wavelengths from the first half of the protocol had no prior dark adaptation phase as they are presented right after the first run. The 90-second DAP pupilloconstriction curve was closest to this second run. Therefore, we chose this period for the main experiments.

We also performed preliminary measurements with the *Up* protocol on ten subjects before the main experiments, to explore the behavior of pupilloconstriction in our setup. In those subjects, we found differences dependent on sex, particularly at the longer wavelengths (not shown). This observation was our rationale to explore this relationship in the actual experiment.

**Experiment I.** *Experiment I* was divided into two protocols, in each of which the stimuli consisted of a stepwise sweep across the available spectrum (Fig 2). One such sweep (*Up*) went from the lowest to the highest wavelength, the other (*Down*) from the highest to the lowest. Light onset happened after a dark adaptation period of 90 seconds duration at the beginning of each protocol (see above), and light remained on during the entire protocol. Light steps lasted 15 seconds each, followed by the next light step at a spectral distance of about 5 nm. We



**Fig 2. Experimental protocols.** (A) Series of peak wavelengths for each experiment and protocol. A list for the respective order of wavelength peaks is available as part of the [S3 File](#). (B, C, D) Schematics of the first 40 seconds of each experimental protocol. Light onset after dark adaptation occurs at zero seconds (dotted black line). Traces show the average pupil diameter, with a ribbon representing the SEM. For visual reasons, only one of the two protocols sharing the same procedure is shown. These are *Up* (B), *Central 1a* (C), and *Short 1* and *Long 1* (D).

<https://doi.org/10.1371/journal.pone.0253030.g002>

hypothesized that with continuous light that shifted slightly in peak wavelength with each step we would see the ipRGC influence in the pupil constriction amplitude by wavelength. Due to the sluggish nature of ipRGC responses [12–14], ipRGC-driven constriction might trail behind, which was an additional reason to balance the spectral protocol series by an up-down alternation. McDougal and Gamlin [13] report that ipRGC-influence at 10 to 18 seconds after light onset takes over as the strongest relative influence on pupil constriction, with rods being a close second; cones were found to have, after 10 seconds, about 1.5 to 2 orders of magnitude lower sensitivity. Since each step in *Experiment I* lasts 15 seconds, we expect a one-step delay to ipRGC pupil constriction, which would equal a 5-nm wavelength shift when looking at just one protocol. Since the changes between each wavelength step were small, the pupils quickly adjusted, usually within one or two seconds after each shift, followed by typical pupil oscillations and pupillary escape. Each protocol consisted of 61 light steps, totaling 915 seconds of light (15.25 minutes), or 1005 seconds including dark adaptation (16.75 minutes). Seventy-five participants were enrolled in *Experiment I*. Due to technical difficulties, the *Down* protocol was not available when *Experiment I* started, and was therefore measured consecutively to the *Up* protocol. Due to this sequence, the number of partaking subjects differs between the

protocols, with 57 participants in the *Up* and 23 in the *Down* protocol. Of those taking part in the *Down* protocol, we measured five randomly selected participants in both protocols to check whether there were any systematic differences between the protocol groups, which was not the case. Mixed models were used in the statistical analysis to take these factors into account.

**Experiment II.** The surprise outcome in *Experiment I* was the comparatively high wavelength of 540 nm at which pupil constriction amplitude was at its maximum (see below under *Results* for all outcomes and the corresponding figures). In *Experiment II*, we thus tested whether this was due to our experimental design of small step changes along the spectrum in *Experiment I*. We wanted the series of wavelengths to be continuous, i.e., that wavelength changes would happen near instantaneous, which would be the same as in *Experiment I*. However, the monochromator required considerably more time for the changes of wavelength between steps than in *Experiment I*. We therefore designed three new protocols (Fig 2A and 2C). All three protocols started, after a dark adaptation phase of 90 seconds, at 548 nm, which is halfway on the available range of wavelengths. Both comprise 61 light steps in total. The series are not random, which would have led to an imprecise estimate of wavelength peaks, as we found in early tests. However, the series are designed to be as diverse as possible in terms of wavelength changes, and regarding when in the series specific wavelengths are presented. Except for the first step, the protocols are symmetrical to the midpoint of the series axis. The protocols *Central 1a* and *Central 2* continued with differing spectral-wavelength shifts between steps. In these two protocols, the shutter was closed during the wavelength adjustment. In the longest case, the wavelength change lasted just under 9 seconds, and we set the shutter closure time to that value for all steps. The result, for both protocols, was 15 seconds of light followed by 9 seconds of darkness at each step, totaling 24 seconds per wavelength; the total series had 1464 seconds after light onset and 1554 seconds including dark adaptation, i.e. just under 26 minutes). Since the presence of dark periods between light steps in the two protocols above introduced a new factor to the experimental design, we designed the further protocol *Central 1b*. This third protocol was identical to *Central 1a*, except that the shutter stayed open at all times, i.e., there were no dark phases in between light steps, and the 15 seconds per light step started right after the wavelength adjustment (Fig 2C). As an undesired side effect, participants perceived the adjustment of wavelengths in that series. We reasoned, however, that these brief periods (about 28 ms per 1 nm wavelength) would not influence pupil constriction amplitude 10 to 15 seconds later in a relevant manner. With all three protocols taken together we could test whether the wavelength dependency in the first experiment was due to the sequence of stimuli and whether the interruption of light in the first two protocols in *Experiment II* influenced the wavelength dependency in this experiment. Since changes in wavelength between light steps were larger in *Experiment II* than in *Experiment I*, the pupillary reactions were also stronger. In protocols with discontinuous light, the phasic part of the pupillary light reaction took about five seconds or less before changing to a tonic movement with typical oscillations and pupillary escape.

Ten participants took part in *Experiment II*. The sample size was chosen based on the first experiment. There, we saw that a random sample of ten out of the available sample for every protocol resulted in the same primary outcome for the spectral dependency. Each participant took part in every protocol.

**Experiment III.** Two main results in *Experiment II* led us to the design of a further experiment. *Experiment II* had shown that the spectral dependency of the nPC depended on whether or not there were periods of darkness between light steps, i.e., whether the light stimulus was discontinuous or continuous. Of particular interest was an apparent shift in wavelength sensitivity over time when light steps were discontinuous, but not otherwise. The experiment had



further allowed looking at the wavelength dependency of the PIPR during the intervals of darkness. We designed *Experiment III* to build on these results. It comprised four protocols (see Fig 2D), each of which consisted of 31 light steps following the 90-second dark adaptation phase. Light steps followed the series structure of *Central 1a* and *Central 2* from *Experiment II*; however, since each light step was longer in *Experiment III*, we halved the number of steps to not overly tire participants. The spectrum was therefore divided into steps of about 10 nm wavelength difference. The series *Long 1* and *Long 2* were designed to analyze wavelength-dependency changes over a more extended period of time. Accordingly, each step's period was extended from 15 to 30 seconds duration, followed by 9 seconds of darkness between the steps (1209 seconds after light onset; 1299 seconds including dark adaptation, or about 21.5 minutes). The protocols *Short 1* and *Short 2* were designed to analyze the PIPR over a more extended period. Accordingly, each step's light stimulus lasted 1 second, with 30 seconds of darkness between steps (in total 961 seconds after light onset; 1051 seconds including dark adaptation, or about 17.5 minutes). In the *Long* protocols of *Experiment III*, pupil behavior was like in *Experiment II*, with an extension to the tonic part due to the prolonged stimulus. Ten participants took part in *Experiment III*. Each participant took part in every protocol.

## Procedure

Participants arrived at the appointed time in the laboratory and were welcomed and seated. Measurements were restricted to daytime hours and were scheduled from 08:00 am to 08:00 pm (taking place, on average, at 01:30 pm  $\pm$  2:45 h:m). The laboratory was a windowless room with constant temperature, mechanical ventilation, and lit with warm-white artificial light (about 50 to 100 Lux, depending on position). Participants read and signed the prepared informed-consent form. Color vision was tested first. Illumination on the tests was cool white, with a high color-rendering index. Participants then filled out a three-part questionnaire, (1) the German translation of the Morningness-Eveningness Questionnaire for testing chronotype (D-MEQ; scores the time of preference; according to the MEQ manual [28], chronotype scores from 60 and above are considered *morning types* or *Larks*, below 40 *evening types* or *Owls*, and in between *Neutral types*); (2) general demographic questions, and (3) general questions regarding participants' current health. The experimenter then fitted the pupillometry apparatus to the participants' head and the participant lied down on a flat treatment couch with a small pillow under the head. The experimenter ensured that the infrared cameras were positioned correctly so that pupil diameter could later be calculated from camera pictures, either by adjusting the distance to the eye (in *Experiment I*) or by presenting a reference scale at eye level (in *Experiment II* and *III*). The Ganzfeld dome was then pivoted over the participants head until the eyes were at a predetermined position and only the dome's inside was visible. Room lighting was then switched off, and a black, opaque curtain between the participant and the experimenter was drawn shut. Thereby the eye-tracker camera output could be monitored on a personal computer without stray light influencing the experimental setup. Participants were instructed to relax and look straight ahead, with minimal blinking. The experimenter repeated the instructions in following procedures if compliance faded. The appropriate protocol was then started, each beginning with its 90-second dark adaptation phase, followed by the respective series (see above). Participants' eyes were not medically dilated, and both eyes received the light stimulus (closed-loop paradigm [19]). If participants took part in more than one protocol (primarily in *Experiment II* and *III*, see above), protocol order was randomized. Participants further got up from the treatment couch for about five minutes in-between protocols during which room lighting was switched on, as in the initial setting. At the end of the experimental session, participants were thanked and debriefed. The total time participants

spent on site was between thirty minutes in *Experiment I* and about two hours in *Experiment II* and *III*.

## Data analysis

Bio-signal data analysis covered converting raw pupil data to a time series of constriction amplitudes for each participant and protocol. It further covered spectral calculations to derive photopic, mesopic, and scotopic illuminance values for each narrowband wavelength peak.

**Normalized pupillary constriction.** We used the *R* software (Version 4.0.2) [34] with several packages for data analysis (*anytime*, *cowplot*, *ggplot2*, *ggmisc*, *knitr*, *readxl*, *signal*, *tidyverse*). Raw pupil data were stored in pixel-based units by the measurement software, and exported to CSV files. These values were converted to mm and mm<sup>2</sup> by a factor derived from the *ImageJ* software [35], using screenshots from the eye tracker videos with included reference scales (see above). Blinks and other artefacts were then removed by calculating a *circularity index* (ratio of pupil height to width, and width to height). Pupil values with a *circularity index* smaller than 0.7 were removed [12, 36]. We also removed pupil diameter values outside reasonable thresholds, usually when above 8.5 mm or below 1.5 mm pupil diameter. However, in some cases, a different cut-off was decided upon after visual inspection. A Savitzky-Golay filter [37] with a third-degree polynomial [26, 38] was used for smoothing of the pupil data; filter length was set to 31 data points or about 0.5 seconds. The same filter was also used to calculate the second derivative of the smoothed curve—valleys in the second derivative coincide with the early stages of pupil constriction. The timestamps of these valleys were used to shift the measurement time according to light onset. The shift was monitored visually and, when necessary, manually adjusted. After this correction, time values are negative during dark adaptation, light onset happens at zero seconds, and the experimental lighting conditions occur at positive time values according to the protocol.

In the next step, the normalized pupillary constriction (nPC) was calculated at each time step  $i$ , as

$$nPC_i = \frac{\text{diameter}_{\text{baseline}} - \text{diameter}_i}{\text{diameter}_{\text{baseline}}} * 100\%. \quad (1)$$

We use the term nPC instead of *pupillary constriction amplitude* suggested by Kelbsch et al. [19]. Both are calculated the same way. However, nPC (since it does not refer to an amplitude) might be the more natural term for the prolonged stimulation in our experiments. The normalized pupillary constriction is defined as the decrease of pupil diameter, normalized to a baseline level (generally the baseline before stimulation [19]), thereby controlling for the considerable variation of inter-individual pupil size, as well as for its (smaller) intra-individual variation [16, 19]. The baseline diameter corresponds to 0% constriction, and the (impossible) pupil diameter of 0 mm would correspond to 100% constriction. As an example, a light-adapted diameter of 2 mm from a baseline pupil diameter of 8 mm equals 75% constriction amplitude. As we used a comparatively short dark adaptation period, followed by a long series of light steps, we believe that the pupil diameter at the end of the dark adaptation phase somewhat loses its prominence, as it is mostly unrelated to any particular wavelength peak across the protocols. Therefore, we used the largest one-second-mean pupil diameter (see below) as the pupil baseline, taken from the period between the last seconds of dark adaptation until the end of each protocol. This baseline still takes interindividual differences into account, while intraindividual differences are already taken care of through the continuous measurement over all wavelengths per person and protocol. By definition, nPC can only take positive values here. To check whether our normalization method affected the results in an undesirable way,

we performed an additional primary analysis without normalization in *Experiment I*, and further tried an alternative normalization in *Experiment II* and *III* to the average diameter during the last second of darkness before each light step. Since we had measurements for both eyes, one eye was discarded at the next step; the eye with the highest count of remaining data points was kept unless visual inspection showed abnormalities. The CSV export from the collected steps above contained the pupil diameter (mm),  $nPC_i$  (%), Series number, time since dark adaptation/light onset (s), time since the start of the light step (s), wavelength (nm), baseline diameter (mm), participant code, eye, date, time, and protocol name. Data in the CSV export file were further aggregated in two ways, so besides unaggregated pupil data (i.e., with 60 Hz sampling frequency) the file contained one-second means, and means over specific periods. The specific periods were the last five seconds of each light step, for every protocol except *Short 1* or *Short 2*, in which the sixth second after light offset was used instead [21]. In a final step before the statistical analysis, the pupil data were combined with participant data (sex, age, chronotype), and spectral data (see Fig 1 and below). The scripts to derive the input data for the statistical analysis are available as *R-Markdown* files as part of the, in the S1 File. All data used in the statistical analyses are available as CSV files from the *Open Science Framework* [39]. Before the experiments we tested and calibrated the eye tracker on a reference head and pupil. Camera resolution, combined with the mean camera-to-eye distance of 27.5 mm, led to a resolution of 10 pixel per mm at pupil level. Pupil height or width changes can thus be tracked to about 0.1 mm accuracy or 1.5% nPC, assuming a baseline pupil of just under 7 mm.

**Spectral calculations.** From the spectral irradiance measurements, total irradiance was calculated automatically by the measurement software (see above). Similarly, illuminance values were automatically calculated according to the CIE-1931 standard 2° photopic observer which uses the  $V(\lambda)$  function [40], as well as according to the CIE-1964 standard 10° photopic observer, which uses the  $V_{10}(\lambda)$  function [41]. The five types of alpha-opic equivalent (D65) daylight illuminance were calculated according to the CIE S 026 standard (*CIE-S-026-EDI--Toolbox-vE1.051*) [32]. We also used the *Melanopic-light-sources\_Toolkit\_V13.12*, which was developed and provided to us by Dieter Lang [42]. The toolkit provided illuminance values for the CIE-1951 scotopic observer with  $V'(\lambda)$  [43] spectral weighing in addition to the alpha-opic values. Illuminance for the mesopic observer was calculated according to the German standards DIN 5031-2 [44] and DIN 5031-3 [45] through  $V_{eq}(\lambda)$  weighing functions. The standards were recommended by Adrian [7]. The CIE has since released the standard CIE 191:2010 for mesopic photometry, which is based on the same basic principles but uses a different calculation procedure and notation ( $V_{mes,m}(\lambda)$ ) [46]. In the German standards, illuminance calculations for the mesopic observer use spectral weighing functions  $V_{eq}(\lambda)$  that are intermediary between the photopic  $V_{10}(\lambda)$  and the scotopic  $V'(\lambda)$  function, in addition to intermediary luminous efficacies of radiation. These intermediaries are chosen based on the adapting equivalent luminance  $L_{eq}$  for a 10° observer on a  $\log_{10}$ -based scale between the photopic (100  $\text{cd}/\text{m}^2$ ) and scotopic ( $10^{-5}$   $\text{cd}/\text{m}^2$ ) endpoints defined by the standards. Our calculations were performed using *Microsoft Excel* software. We used the corneal stimulus intensity (log irradiance, without prereceptor filtering) in our analysis. Estimates for prereceptor filtering on irradiance and photon density are available as part of the S2 File. We used the method from Lund et al. [47] to calculate the age-dependent filtering as suggested in the CIE S 026 standard [32]. A visual comparison of spectral irradiance with vs. without prereceptor filtering for the median age of 26 years is displayed in the S13 Fig. Peak wavelength-shifts due to prereceptor filtering were small and did not occur above 428 nm. Spectral irradiance with filtering had a higher peak wavelength of 1 nm for 428, 424, and 420 nm; it was 3 nm higher for 416 nm and 409 nm, and 5 nm higher for 404 nm. These changes in peak wavelength are identical through our age range between 18 and 36 years. The spectral measurements for the range

of peak wavelengths, including total irradiance and all types of illuminance values are provided as part of the *Supporting Information*, in the [S2 File](#).

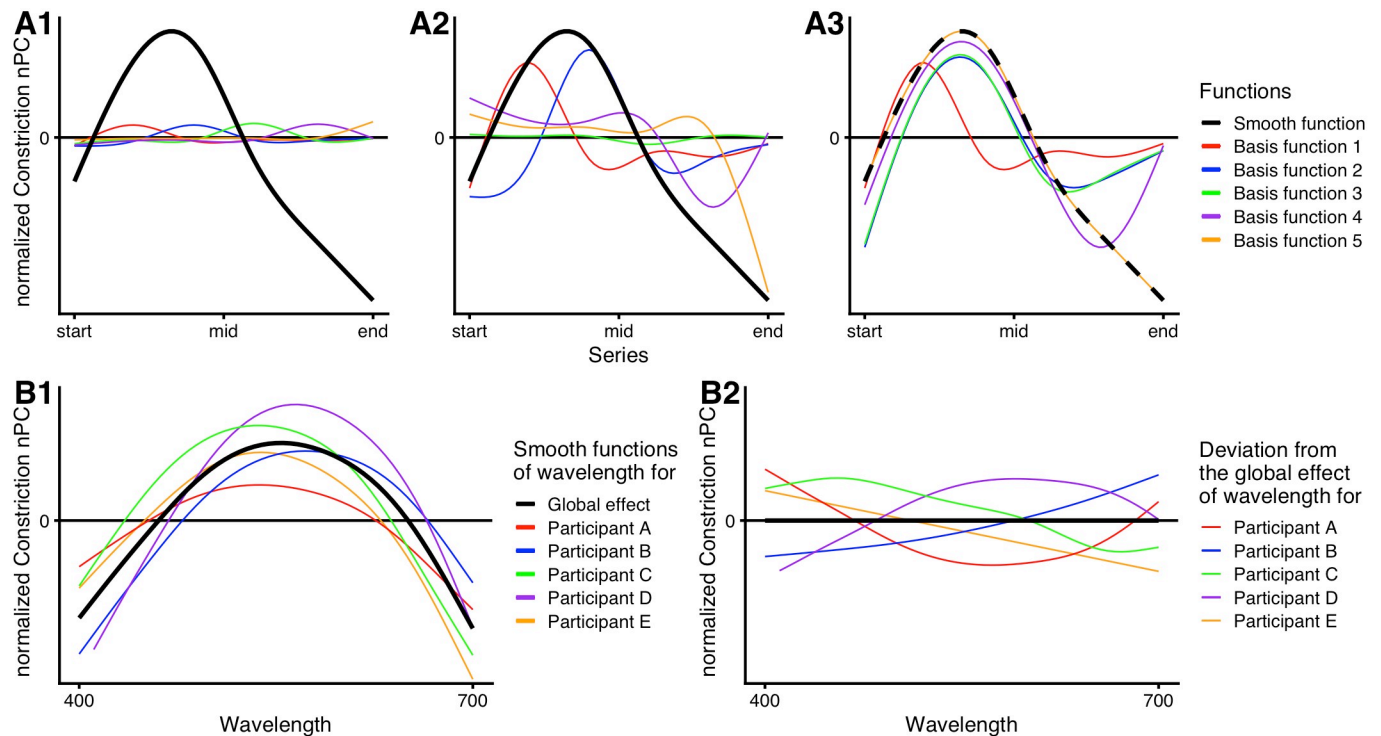
## Statistical methods

We used the *R* software (Version 4.0.2) [34] with the *mgcv* (Version 1.8–31) [48] package to perform a *generalized additive mixed-effect analysis* on the connection between normalized pupillary constriction and the empirically and theoretically derived predictors. *Generalized Additive Models* (GAMs) [48, 49] allow for a data-driven decomposition of the relationship between a dependent variable and user-defined predictor variables in both a parametric and nonparametric fashion. A variant of GAMs are *Generalized Additive Mixed-Models* (GAMMs or HGAMs), used in the context of hierarchical data as is the case in any repeated-measures setup such as ours [50]. GAMs are widely used in *Biology*, *Ecology*, and *Linguistics* [48]. GAM (M)s are not yet used often in *Human Life Sciences* [51], but we find them an excellent match for the analysis at hand, and mixed-effect models gain support in biophysiological research such as visual perception [52] and nonvisual effects of light [53]. Guidelines on the theory behind GAMs and their practical use can be found in Wood [48], Simpson [54], Pedersen et al. [55], and the guide by Sóskuthy [50].

The nonparametric parts of GAMs take the form of *smooth* functions (*smooths*, in short), describing the connection between the outcome variable and a predictor. A *smooth* function  $f(x_i)$  is the weighted sum of a number  $k$  of basis functions  $b_j$  at the  $i$ -th value of a predictor variable  $x$ :

$$f(x_i) = \sum_{j=1}^k b_j(x_i)\beta_j \quad (2)$$

The  $\beta_j$  in [Eq 2](#) denote the weights (the *smooth* coefficients) of the basis functions  $b_j(x_i)$ . Since GAMs are not yet widely familiar in this field of research, their concept is visualized in [Fig 3A](#). The type of basis function is chosen as part of the model construction, with *cubic regression splines* being probably the most well-known type ([Fig 3A1](#)). Unless otherwise stated, we used the *mgcv*-default type of basis functions, which are *thin plate regression splines* (TPS). Wood [48] gives a comprehensive overview of basis functions and their suitability depending on the context. Unlike coefficients in a normal parametric regression, the basis-function weights  $\beta_j$  in generalized additive models are not by themselves interpretable towards the modelled relationship ([Fig 3A2](#)). Only by knowledge of the set of basis functions and their eponymous weighted addition (“*additive*” models) can the resulting *smooth* function be interpreted ([Fig 3A3](#)). One of the underlying assumptions to the *smooths* is a constancy of complexity across the value range of  $x$ . That means that *smooths* do not do well without additional adjustments, if constructed from relationships that vary heavily in their dynamics across the range, or that exhibit discontinuous changes (step-changes). These adjustments include the use of *adaptive splines* for varying dynamics, or the use of factors for step-changes [48]. *Smooth* functions can be multi-dimensional, thereby describing interdependent relationships or interaction effects between predictors and the outcome variable. The number  $k$  of basis functions is chosen based on the complexity of the modelled relationship present in the data. Overfitting is avoided through the penalization of complexity as part of the likelihood optimization. Therefore, added complexity (or *wiggleness* [48]) in the *smooth* needs to improve the model fit enough—compared to, e.g., a straight line—to outweigh the penalization. More broadly, GAMs ascertain to model the simplest possible relationship between variables, or no relationship at all, without leaving out relevant structural components depending on any given predictor.



**Fig 3. Concepts of additive (mixed-effect) models.** (A) Construction of a *smooth* through *basis expansion*. Every *smooth* is constructed from a number of basis functions, usually spread evenly across the value range of a predictor. In the shown case, five cubic regression splines were used as basis functions to demonstrate the concept. The thick black line represents the final *smooth* function, describing the relationship between the predictor and the outcome variable. Colored traces show the basis functions. (A1) Unscaled basis functions—their type and maximum number are part of the input for model generation. (A2) Scaled basis functions—for each basis function, a respective weight is estimated by which the function is scaled. (A3) Summation of basis functions, starting with the first and then, successively, adding the others; the generated *smooth* can then be used for prediction. The *smooth* is shown as dashed curve to show that, with the addition of the last basis function (orange line), the resulting curve is equal to the *smooth* function. (B) Concept of global effects with random *smooths*. (B1) The thick black line represents the global effect, describing the average relationship between the predictor and outcome variable. Colored traces show the individual's effect of the same predictor, demonstrating interindividual differences. The model takes these differences into account through the so-called random *smooths* (i.e. “random” in that their contribution depends on the subject). (B2) Random *smooths* are *smooths* describing the deviations from the global effect. The colored traces show the deviations present in panel B1. Because the global *smooth* and the random *smooths* are estimated together, the deviations disappear on average, i.e., not all deviations will tend in the same direction, but are spread around the global effect.

<https://doi.org/10.1371/journal.pone.0253030.g003>

If not stated otherwise, we used *Akaike's Information Criterion* (AIC) for model selection, as suggested by several sources on GAM(M)s [48, 50, 54–56]. The AIC incorporates model fit (likelihood) and model size (number of parameters) [48]. It can further be compared between similar models, with the lower value indicating the better-suited model. Following the example of Pedersen et al. [55], models that differ by two units or less from the lower AIC have substantial support, with the more parsimonious model to be preferred. We report  $\Delta$ AIC values when there is support for the inclusion of a predictor. T-tests obtain p-values for parametric terms. The p-values for nonparametric terms are *approximate p-values*, so called owing to the complexity of the *degrees-of-freedom* (*df*) concept used for statistical testing, and other underlying assumptions [48, 50]. For nonparametric terms, the hypothesis tested-against states that the relationship between predictor and dependent variable is a horizontal, flat line. P-values less than or equal to 0.05 were considered significant. Confidence intervals in prediction plots account for the model's overall uncertainty, not for the respective plotted predictor alone.

We explored the nPC's relationship with *wavelength*, *irradiance*, *series*, *time*, *time of day*, *sex*, *chronotype*, and *age* by the methods described above. We further looked for the nPC's best



fit to *photopic, scotopic, and mesopic illuminance*, and *alpha-opic equivalent-daylight illuminance* values. The variable *wavelength* refers to the dominant or peak wavelength of the narrowband stimuli as described above. After light offset, *wavelength* refers to the prior light stimulus. *Series* refers to the numbers of light steps after the dark adaptation phase. *Time* refers to the time since the start of the current light stimulus, i.e., light onset or change in wavelength. Several additional R packages were used for the statistical analysis and plot generation (*tidyverse*, *ggplot2*, *readxl*, *dplyr*, *cowplot*, *lubridate*, *itsadug*, *printr*, *patchwork*, *here*, *reshape2*, *plotly*, *gganimate*, *gridGraphics*, *transformr*, *glue*, *magick*, and *DT*). All scripts for analysis and plot generation are available as R Markdown files in the *Supporting Information, S3 File*.

**Base model structure.** While the relevant code is included as part of the *S3 File*, we believe it helps explaining the basic model structure and settings used to analyze the experimental outcome. The model structure is one of five basic GAMM variants (according to Pedersen et al. [55]), which differ in how fixed and random effects are included. The chosen variant, by way of AIC comparison and model diagnostics, is shown in the following equation:

$$E(nPC_i) = \alpha + f_1(\text{wavelength}_i) + f_2(\text{series}_i) + \beta_1 \times \log_{10}(\text{irradiance}_i) + b_p + f_p(\text{wavelength}_i),$$

$$\text{where } nPC_i \sim \frac{N(\mu_i, \sigma^2)}{\sqrt{\frac{\chi^2}{n}}}, \text{ and } b_p \sim N(0, \sigma_b^2) \quad (3)$$

Eq 3 shows the statistical base model, where the expected value of nPC  $E(nPC_i)$  is modelled by *smooths* as a function of *wavelength* ( $f_1$ ) and *series* ( $f_2$ ), and a parametric effect of  $\log_{10}(\text{irradiance})$ ;  $\alpha$  is the average nPC, when all other terms are zero, also called the intercept;  $\beta_1$  is the parametric coefficient, or slope, for  $\log_{10}(\text{irradiance})$ ;  $b_p$  are random intercepts by participant; and  $f_p(\text{wavelength}_i)$  are random *smooths* by participant.

*Wavelength* is the main predictor, *series* was included for empirical reasons, i.e., in the cases when visualization of the nPC data in *Experiment I* showed the influence of the protocol (*series*). As stated above under *Apparatus*, *irradiance* is theoretically motivated to compensate for lower irradiance levels under 450 nm wavelength. However, *smooth* behavior from 400 to 450 nm still has to be interpreted tentatively, since the estimate for *irradiance* might under- or overcompensate the effect, thereby distorting the influence of the lowest wavelengths on nPC. We account for this by leaving *irradiance* out of the model in a secondary analysis and discuss differences between the variants.

As random effects in *Experiment I* we had random intercepts  $b_p$  and random *smooths*  $f_p(\text{wavelength}_i)$  for each participant  $p$ . In *Experiment II* and *III*, all participants took part in all protocols; we, therefore, had random effects for participants-by-protocol in those cases [50]. Random intercepts describe the deviation from the average nPC ( $\alpha$ ) depending on the participant (or participant-and-protocol), with independent and gaussian distributed values. Random *smooths* can be thought of as the nonparametric version of random slopes in *linear mixed-effect models*. Random *smooths* allow for the individual's deviation from the global effect  $f_1(\text{wavelength}_i)$ , but zero-out on the global effect when viewed on average across participants (Fig 3B).

For the distribution of our response variable,  $PCA_i$ , which takes positive values between 0 and 1, the *gamma* or *beta* distribution seemed sensible assumptions initially. However, model diagnostics with those were poor compared to even the standard *gaussian* error distribution. Instead, we thus went with the *scaled-t* (*scat* in *mgcv*) distribution family, which is less susceptible to outliers and indeed greatly improved residual distribution. Eq 3 states that  $PCA_i$  varies around its mean  $\mu_i$  with a t-distribution based on the standard deviation  $\sigma$  and  $n$  degrees of freedom. We also tested a model with the *gaussian location-scale* (*gaulss* in *mgcv*) family. This distribution allows the variance to change depending on a predictor, just like the mean.

However, we did not find relevant differences in variance, and the model with the *scat* family performed similarly well with respect to the AIC. The *gaulss* family did further not allow for the computationally faster *bam* command, so we used the *scat* family for the analysis.

We used the *gam.check* function of *mgcv*, which includes several residual plots for model diagnosis. Visual inspection of these plots did not reveal any apparent deviation from homoscedasticity or normality. Models were further controlled for autocorrelation in the residuals. In all models that used only one timestep per wavelength peak (5-second average nPC), the inclusion of random effects diminished all relevant autocorrelation. In models that used 1-second averages for nPC, random *smooths* by *wavelength* left serious residual autocorrelation present. Random *smooths* by both *wavelength* and *time* proved too computationally intensive in those cases and we instead included the remaining autocorrelation with an autoregressive error model [50, 54], in addition to random *smooths* by *wavelength*. Finally, the number of knots for each *smooth* (i.e., how many basis functions comprise the *smooth*) was also checked (*k.check*) and was increased, when theoretically sensible, until autocorrelation along the *smooth* vanished [55]. For computational speed reasons, we used the *bam* function with *fREML* in R instead of the standard function of *gam* with *REML*. *Bam* is optimized for big data sets and model structures.

If not otherwise stated, GAMs in the *Results* section were used to model normalized pupillary constriction (nPC) according to Eq 3, the latter also called the base model. Additional parametric and nonparametric predictors were explored—their addition to or removal from Eq 3 are stated in the respective sections. The default nPC is the average value of each light step's last five seconds, as described above.

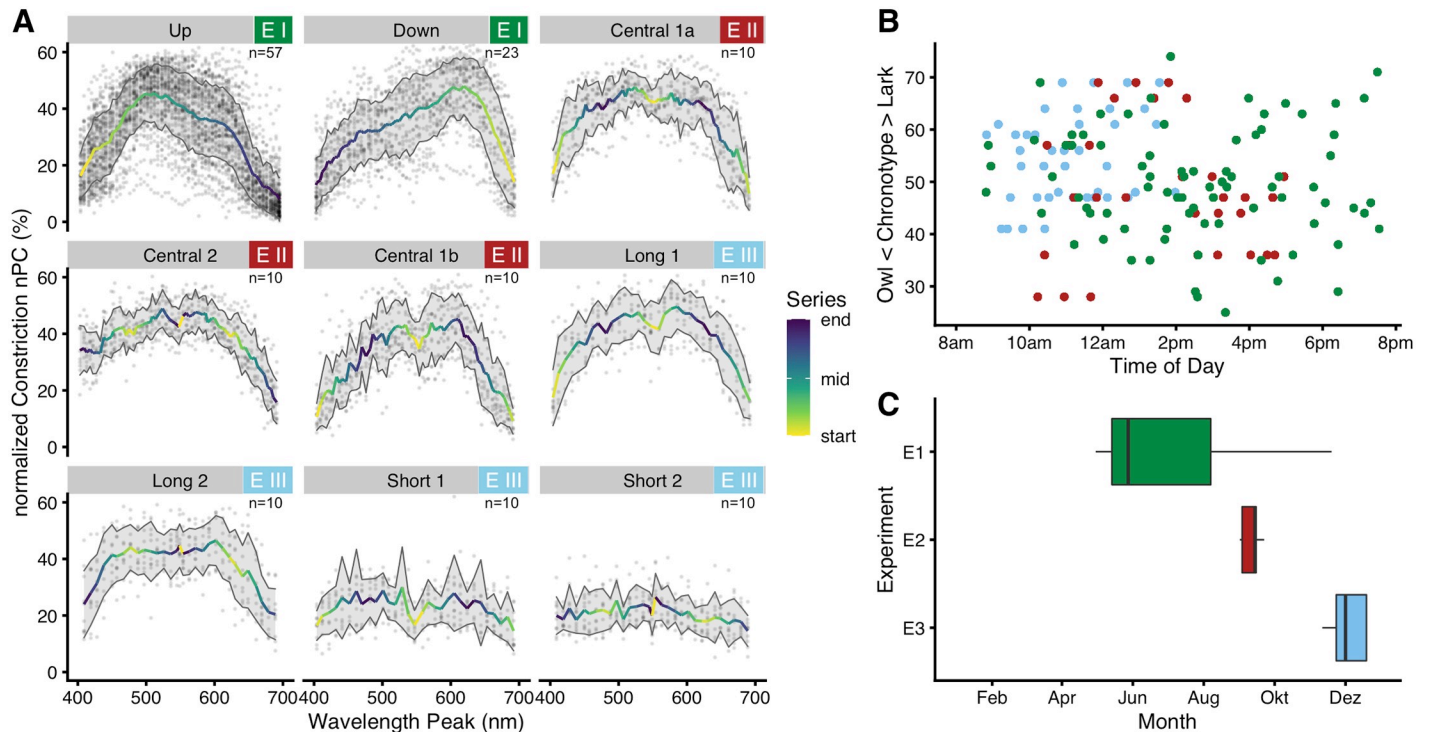
**Linear mixed-effect model structure.** We used the *lme4* package [57] to perform a *linear mixed-effects analysis* of the relationship between nPC and the various illuminance variables described above:

$$E(nPC_i) = \alpha + (\beta_1 + \beta_p) \times \log_{10}(\text{illuminance}_i) + b_p, \\ \text{where } nPC_i \sim (\mu_i, \sigma^2), b_p \sim N(0, \sigma_b^2), \text{ and } \beta_p \sim N(0, \sigma_\beta^2) \quad (4)$$

Intercept  $\alpha$  and slope  $\beta_1$  are the fixed effects of the model described in Eq 4. The intercept  $\alpha$  indicates the value of nPC, when all other terms in the equation are zero. The slopes, or beta coefficients,  $\beta_1$  and  $\beta_p$ , represent the change in the expected value of  $nPC_i$  when increasing illuminance by one log unit. As random effects we included random intercepts by participant,  $b_p$ , and random slopes by participant,  $\beta_p$ , both of which have independent values with a *gaussian* distribution. Random intercepts  $b_p$  show how much individuals deviate from the average nPC  $\alpha$  with a variance of  $\sigma_b^2$ . Random slopes, similarly, show how much the relationship between the predictor and outcome variable deviate from the fixed effect  $\beta_1$  with the variance  $\sigma_\beta^2$ .  $nPC_i$  has a *gaussian* distribution, with variance  $\sigma^2$  around its mean  $\mu_i$ . The proportion of variance-accounted-for ( $R^2$ ) was calculated according to Xu [58].

## Results

An overview of the experimental data and circumstantial conditions is shown in Fig 4. Fig 4A shows plots for nPC vs. wavelength for each protocol, where the underlying data are not controlled for any dependencies. Non-normalized pupil diameters vs. wavelength are shown in S12 Fig. Fig 4B shows *chronotype* vs. *time of day*. Noteworthy is the lack of *Owl* chronotypes before midday, as well as pronounced *Larks* in the morning. Fig 4C shows at what time in the year the experiments took place, as suggested by Veitch and Knoop [59]. Experiments took place in 2019 (preliminary measurements, *Experiment I*) and in 2020 (all experiments).

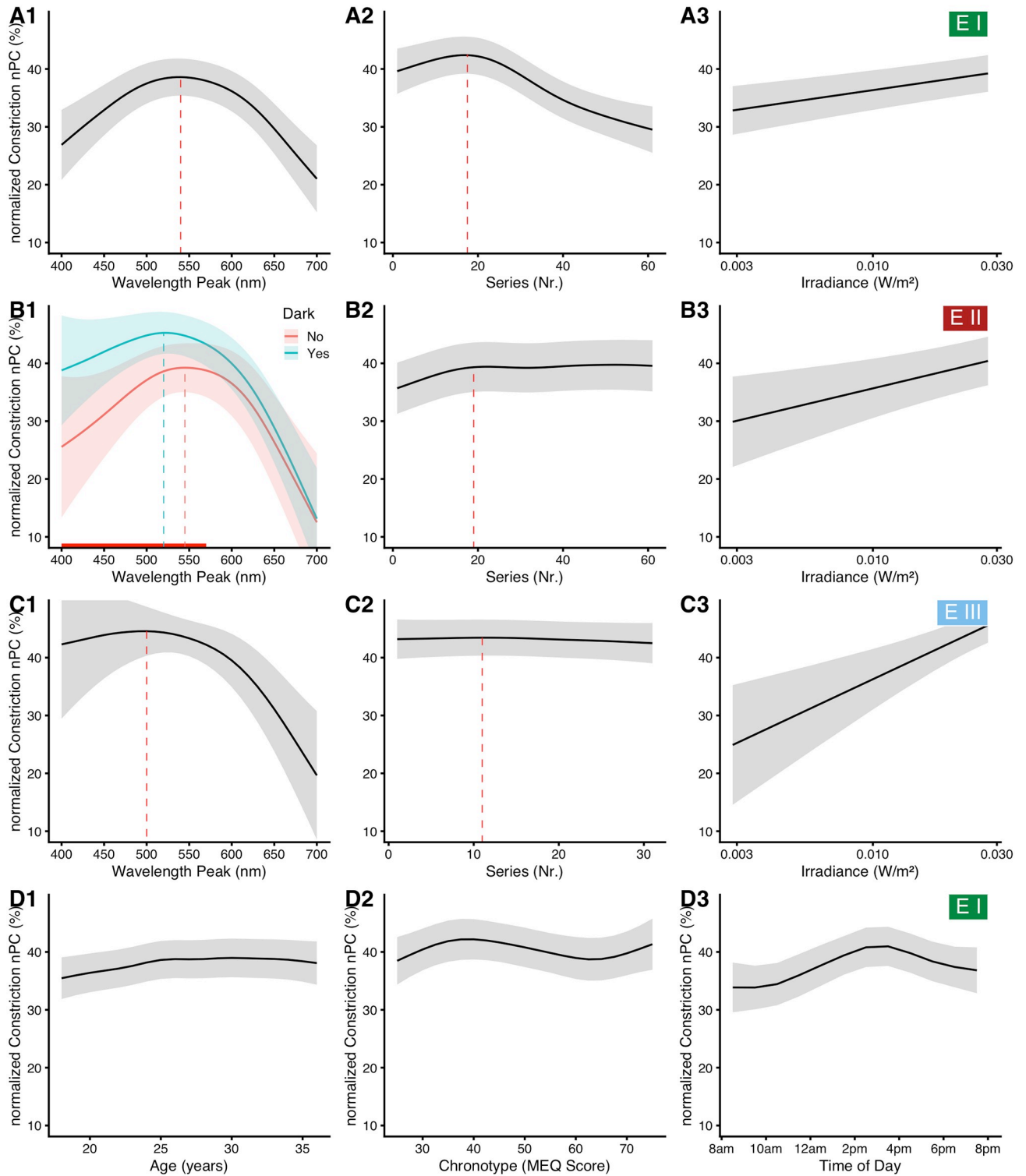


**Fig 4. Experimental data and circumstances.** (A) Normalized pupillary constriction (nPC) plotted against wavelength, for each of the nine protocols. The color scale shows at which point in the series a specific wavelength was presented; light yellow represents early in the series. For all but the *Short* protocols, points represent the average nPC during the respective last five seconds of a light step. In the two *Short* protocols, points represent the average nPC during the sixth second after lights-off (or seventh second after lights-on). Traces show the mean nPC, ribbons its standard deviation. The number in the upper right corner of each plot shows the corresponding sample size. (B) Scatterplot of when subjects of a certain chronotype started their respective protocols. The color scheme is according to (C), i.e. green, red, and blue correspond to Exp. I–III, respectively. Note the lack of *Lark* and *Owl* chronotypes in *Experiment I* before midday. (C) Boxplot of the time of the year when the experiments took place. Preliminary measurements and parts of *Experiment I* took place in 2019, the other measurements in 2020.

<https://doi.org/10.1371/journal.pone.0253030.g004>

## Experiment I

**Base model results.** The results for the base model of all three experiments are shown together in Fig 5, for better comparison. The results for *Experiment I* are shown in Fig 5A and are reported in this section; the results for *Experiment II* and *III*, as well as secondary results for *Experiment I* are shown here in Fig 5B–5D, respectively, but are reported below in the respective sections. The statistical base model contains *wavelength* and *series* as nonparametric predictors, and  $\log_{10}(\text{irradiance})$  for the average nPC during the last five seconds of each light step as a parametric predictor. There is strong support for a dependence on *wavelength* in its implemented form versus other random effect structures (minimal  $\Delta\text{AIC} = 33$ ), or no wavelength dependence ( $\Delta\text{AIC} = 4963$ ), as there is for including *series* as predictor ( $\Delta\text{AIC} = 390$ ). *Irradiance* did improve the model further ( $\Delta\text{AIC} = 17$ ). All p-values in the final model were below 0.001. The dependence on *wavelength* shows an inverted-U shaped curve with a peak at 540 nm, and an average nPC of about 38% when controlling for all other factors (Fig 5A1). Compared to the maximum, nPC is predicted by the model as up to 11% lower for short wavelengths and up to 17% lower for long wavelengths. The dependence on *wavelength* shows a peak at 550 nm when not controlling for *irradiance* (S11A Fig). With respect to series effects, nPC increases slightly with *series* up to step 18 (by about 3%), which occurs at about 360 seconds after the protocol start, or, respectively, 270 seconds after lights-on (Fig 5A2). After this point, nPC declines by about 13% until the end of the protocol. nPC further increases with



**Fig 5. Model predictions.** Model predictions for normalized pupillary constriction (nPC) as depending on several main predictors, when all other predictors are held at an average, constant level. Traces show the model prediction for the mean, ribbons its 95% confidence interval. Dotted lines show a peak that is discussed in the main text. In (A3, B3, and C3), the x-axis is logarithmically scaled to reflect the logarithmic transformation of irradiance in the model. (A) and (D) show dependencies in *Experiment I*, discussed here, (B) and (C) dependencies in *Experiment II* and III, respectively, discussed later. (B) The red line at the bottom of (B1) indicates where the two curves differ significantly at the 0.05 level of significance.

<https://doi.org/10.1371/journal.pone.0253030.g005>

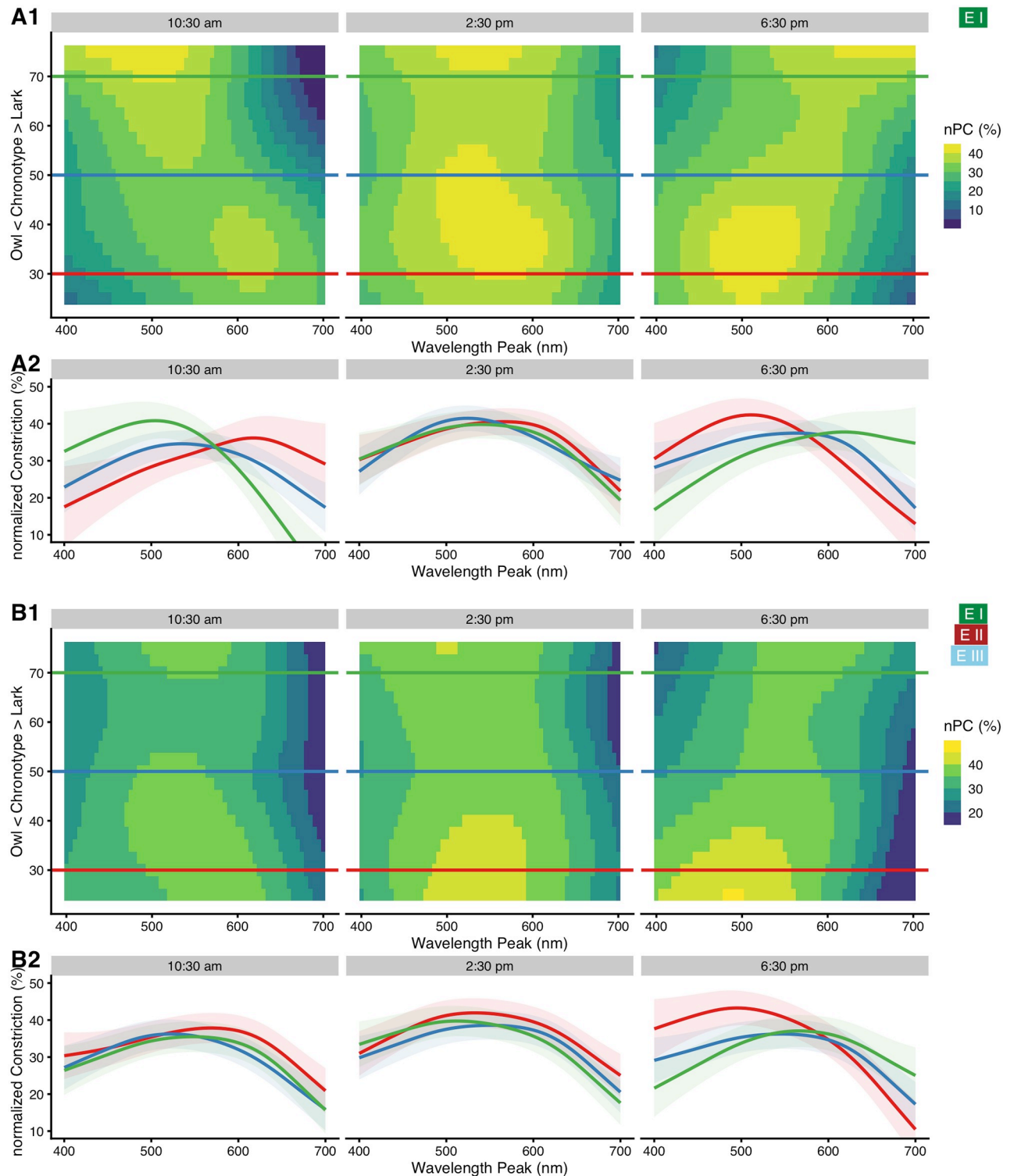
*irradiance* ( $\beta_{\log_{10}(\text{irradiance})} = +6.5\% \pm 1.4\% \text{ SE}$ ), as seen in Fig 5A3. The standard deviation of the random intercept by participant,  $\sigma_b^2$ , was 8% in *Experiment I*. While a balanced sample between the *Up* and *Down* protocols would have been preferable, GAM(M)s allow for imbalanced data. Larger sample groups are not weighted more heavily, and global effects are adjusted for across groups. We also performed the base-model analysis with a random sample of ten participants from the *Up* and *Down* protocols, obtaining the same outcome for wavelength dependency (see the S3 File). We thus decided on a sample size of ten subjects for *Experiment II* and *III*.

**Interaction of series and wavelength.** There is some support for an interaction of *wavelength* with *series*, i.e., that changes of nPC with *wavelength* depend on when, in the series, specific wavelengths are presented ( $\Delta\text{AIC} = 5$ ). Results for that are shown in a (S2 Fig). The model prediction with the interaction effect shows little or no deviation from the basic model results at the instances with data. Compared to the model without interaction, nPC is predicted to be lower for very short and long wavelengths occurring right after light onset. Wavelengths around 600 nm further lead to a more pronounced nPC during the early steps of the series, while for wavelengths around 500 nm nPC is slightly lower during the later steps. In terms of ipRGC sensitivity, it seems as though, compared to the model without the interaction, nPC is heightened at, and a few steps after, reaching the sensitive wavelengths around 490 nm, but the effect is small. With the interaction effect as part of the model, *irradiance* is no longer significant ( $p = 0.31$ ).

**Sex and age.** There is further support for an influence of *sex* on the nPC ( $\Delta\text{AIC} = 85$ ). On average, women had a smaller nPC than men ( $\beta_{\text{women}} = -4.4\% \pm 0.4\% \text{ SE}$ ,  $p < 0.001$ ). The mean difference in pupil constriction (not normalized) between the groups was  $0.21 \text{ mm} \pm 0.024 \text{ mm SE}$ . Our preliminary measurements indicated that the *wavelength* dependency of nPC is conditional on *sex* (not shown). There was, however, no interaction of *sex* with either *wavelength* ( $p = 0.98$ ) or *series* ( $p = 0.27$ ) based on the larger sample from *Experiment I*. With *age* as a predictor (Fig 5D1), nPC increased by about 3% from age 18 up to about age 25 ( $\Delta\text{AIC} = 245$ ,  $p < 0.001$ ). At higher age, nPC reached a plateau. We did not see changes in *wavelength* with *age* ( $p = 0.67$ ).

**Chronotype and time of day.** There is strong support to include *chronotype* and *time of day* as predictors, together with a three-way interaction of these with *wavelength* ( $\Delta\text{AIC} = 47$  and  $\Delta\text{AIC} = 130$ , respectively). Main effects of *chronotype* and *time of day* are shown above in Fig 5D2 and 5D3. In general, *Owls* have a slightly higher nPC compared to *Larks* (Fig 5D2), and nPC was highest during the early afternoon around 3 pm (Fig 5D3). According to the interaction model, subjects with neutral chronotype (chronotype score of 50) showed no substantial change in wavelength dependency across the day (Fig 6A2, blue curves). In contrast, subjects with pronounced chronotype did (Fig 6A2, green and red curves; for a complete overview from 8:30 am to 7:30 pm see S4 Fig). *Larks* (green traces) were shifted in sensitivity towards short wavelengths before midday, were about equal to *Neutral* types at around noon and early afternoon, and were more sensitive to longer wavelengths in the late afternoon. *Owls* (red traces), on the other hand, were shifted opposite; before midday they were more sensitive to longer wavelengths, were close to *Neutral* types at midday and early afternoon, and were more sensitive towards shorter wavelengths during the evening hours. Not every point of this three-dimensional interaction structure has enough data points to support all predictor variable combinations. Missing combinations occur in particular for extreme chronotypes and measurement hours, as was shown in Fig 4B above. However, enough structural points are present to support the dependency thus described (see S5 Fig for more details). To test whether the chronotype effect that we saw was confounded by *age* or by a covariance between





**Fig 6. Interaction of wavelength, chronotype, and time of day.** (A1) False-color graph of model predictions for the nPC's dependence on wavelength (x-axis) and chronotype (y-axis) for three times of day, when all other predictors (basic model) are held constant at their average. Horizontal lines show where the respective three traces shown in figure part (A2) are taken from. (A2) Model predictions for the nPC vs. wavelength for three chronotypes (trace color), and three times of day. Green traces show *Larks* (CT score = 70), red traces *Owls* (CT score = 30), and blue traces *Neutral* types (CT score = 50). Ribbons show the 95% confidence interval for the predicted means. (B1/B2) Like (A), but for pooled data across all experiments. See the main text for further details.

<https://doi.org/10.1371/journal.pone.0253030.g006>

*chronotype* and *time of day*, we calculated the Pearson correlation between the variables. Neither of the combinations of *age*, *chronotype*, or *time of day* had a strong correlation (all  $r \leq 0.13$ ) or was even significant (all  $p \geq 0.26$ ). To increase the number of combinations, we pooled the data from all experiments. The results from the pooled data model are shown in Fig 6B, but reported further below.

**Other results.** When all significant terms are added to the model—i.e., *the interaction of wavelength and series*, the main effects of *sex* and *age*, and further the *interaction of chronotype with time of day* and *wavelength*—all model terms remain significant (all  $p < 0.001$ ) except for *irradiance* ( $p = 0.31$ ). More importantly, the predictor-response relationship across the terms is very similar to the descriptions above. One exception is the effect of *age*. When controlling for *chronotype* and *time of day*, nPC is predicted to gradually rise across the whole range from age 18 to 39 (S3B Fig), as opposed to only the range from age 18 to 25 (S3A Fig). This full model was also the best in terms of AIC yet (min  $\Delta\text{AIC} = 58$ ). AIC would have been further improved by dropping the main effects of *chronotype* ( $\Delta\text{AIC} = 7$ ) or *time of day* ( $\Delta\text{AIC} = 40$ ). However, it is not advisable to drop the main effects in the presence of an interaction [48]. Curiously, most of the random *wavelength*-by-participant *smooths* were not significant in this model (61 out of 75). This indicates that the fixed effects above account sufficiently for the interindividual differences in *wavelength* dependency from roughly 80% of participants.

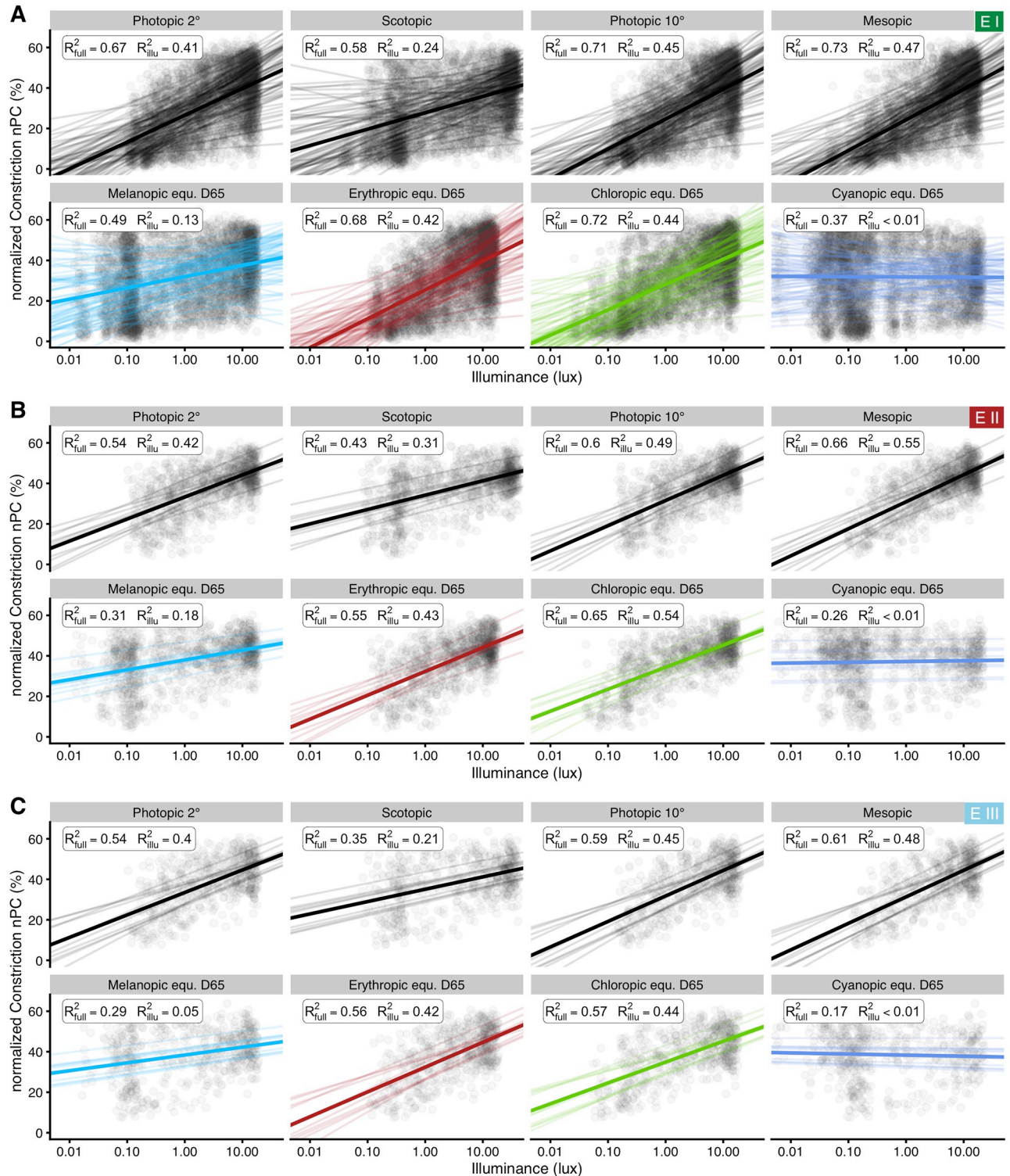
nPC was calculated for the centered and scaled pupil size, as described above in *Materials and methods* (Eq 1). When the pupil diameter is not scaled but is centered on the baseline, results show the pupillary constriction in mm instead of in per cent. When pupil diameter is neither scaled nor centered, results show the raw pupil diameter in mm. We explored those two variants of the response variable, but results do not indicate that such scaling changes the model's composition or interpretation (S6 Fig).

**Illuminance.** To analyze how well each type of illuminance can predict nPC, we constructed several simplified standard linear mixed-effect models according to Eq 4, with a fixed effect for illuminance, and random intercepts and slopes per participant. Illuminance values were  $\log_{10}$  transformed. Nine measures of illuminance were used as described above in *Materials and methods*: photopic illuminance for foveal and ganzfeld stimulation ( $2^\circ$  and  $10^\circ$  observer according to CIE standards), mesopic and scotopic illuminance, and the *alpha-opically weighted equivalent daylight illuminances* for the five types of photoreceptor. Since the scotopic illuminance and rhodopic equivalent daylight illuminance use the same underlying action spectrum, results on model fitting are identical between these two types of illuminance. Therefore, only results for the scotopic model will be shown in the figures. Fig 7 shows the results for all three experiments together for better comparison. The results for *Experiment I* (Fig 7A) show that the photopic ( $2^\circ$  and  $10^\circ$ ) and especially mesopic functions are much better at predicting nPC ( $R^2_{\text{illu}} \geq 0.41$ ) than the commonly used melanopic and rhodopic (viz. scotopic) receptor functions ( $R^2_{\text{illu}} \leq 0.24$ ). Of the three cone functions, erythropic and especially chloropic weighing led to good results ( $R^2_{\text{illu}} \geq 0.42$ ), whereas cyanopic weighing was the only illuminance type with no support at all by the model ( $p = 0.37$ ,  $R^2_{\text{illu}} < 0.01$ ).

Results for *Experiment II* and *III* are shown in Fig 7B and 7C, respectively, for better visual comparison across the experiments. Their results are reported below in the respective sections.

## Experiment II

**Base model results.** The results of *Experiment II* for the base model were shown above in Fig 5B but are described here. The base model is constructed identically to that in *Experiment I*, with an added factorial predictor for whether there is, or is not, a period of darkness between light steps (referred to as *Dark* in the remainder). There is support for including this factor



**Fig 7. Linear mixed-effect model results for nPC's dependency on illuminance.** Model predictions for nPC vs. various measures of illuminance. Points show individual data. Thick regression lines show the fixed-effect relationship and thin regression lines random effect variation in slope and intercept by participant. The insets show  $R^2_{full}$  for the full model (fixed and random effects), and, more importantly, a partial  $R^2_{illu}$ , i.e. the proportion of variance explained through nPC's relationship with the fixed effect of illuminance. Part (A) shows results for *Experiment I*, (B) for protocols with darkness between light steps in *Experiment II*, and (C) for protocols with thirty seconds of light, followed by nine seconds of darkness in *Experiment III*.

<https://doi.org/10.1371/journal.pone.0253030.g007>

( $\Delta\text{AIC} = 5$ ). The dependence on *wavelength* shows an inverse-U shaped curve in both cases (Fig 5B1). The difference between the two curves is about 13% nPC in the short-wavelength spectrum but almost disappears towards the long wavelengths, above about 600 nm (it is significant below 570 nm at the 5% level). nPC increases slightly with *series* up to step 19 (about 3%) and is steady afterwards (Fig 5B2). nPC further increases with *irradiance* ( $\beta_{\log_{10}(\text{irradiance})} = +10.7\% \pm 3.4 \text{ SE}$ , Fig 5B3). When excluding *irradiance* as a predictor, the shorter wavelengths become less influential; the difference between the curves remains similar (S11B Fig). The standard deviation of the random intercept by participant,  $\sigma_b^2$ , was 5.3% in *Experiment II*.

**Time-course of the wavelength dependency.** As stated above in *Materials and methods*, we aggregated the 60Hz-resolution data to mean values over specific periods, i.e., the last five seconds of light, but also created mean values over each one-second period. We used these one-second values to explore how nPC vs. *wavelength* changed over the time course after light onset for each wavelength (referred to as *time*). The results are shown in Fig 8A1 and 8A2 (shown in full in the S7 Fig and S4 File). Note that this implies an oversimplification for the first one or two seconds, where great changes in pupil diameter occur very quickly (cf. Fig 2B, 2C and 2D); the purpose here is to get an overview for the later stages of every light step. nPC decreased over the time course (vertical direction in Fig 8A1, horizontal sequence of graphs in Fig 8A2). Short wavelengths are especially influential during the first seconds of light onset. Peak nPC also shifts slightly towards lower wavelengths between five and fifteen seconds after onset in the case with darkness between light steps (dashed curve, shift from 525 nm to 500 nm, see also S7 Fig), but not otherwise (maximum is stationary at 550 nm). nPC increases with *irradiance* ( $\beta_{\log_{10}(\text{irradiance})} = +9.0\% \pm 1.9 \text{ SE}$ ). When leaving *irradiance* out of the model, differences between the two settings of *Dark* stay mostly the same, whereas nPC sensitivity overall shifts slightly towards higher wavelengths (about 10 to 20 nm; S8 Fig). Finally, settings with darkness between light steps allow for nPC analysis during the re-dilation phase, known as post-illumination pupil reflex (PIPR, Fig 8A2 *Dark*, and Fig 8A2 “21<sup>st</sup> second or 6<sup>th</sup> second after lights-off”). Interestingly, re-dilation is slowest for the shortest wavelengths, which is also the case when not controlling for *irradiance* (S8 Fig). Wavelengths above 600 nm were least influential on the PIPR (Fig 8A2 “21<sup>st</sup> second or 6<sup>th</sup> second after lights-off”). The results for the protocols Long 1 and Long 2 of *Experiment III* are shown in Fig 8B for better comparison with *Experiment II*, but are reported further below.

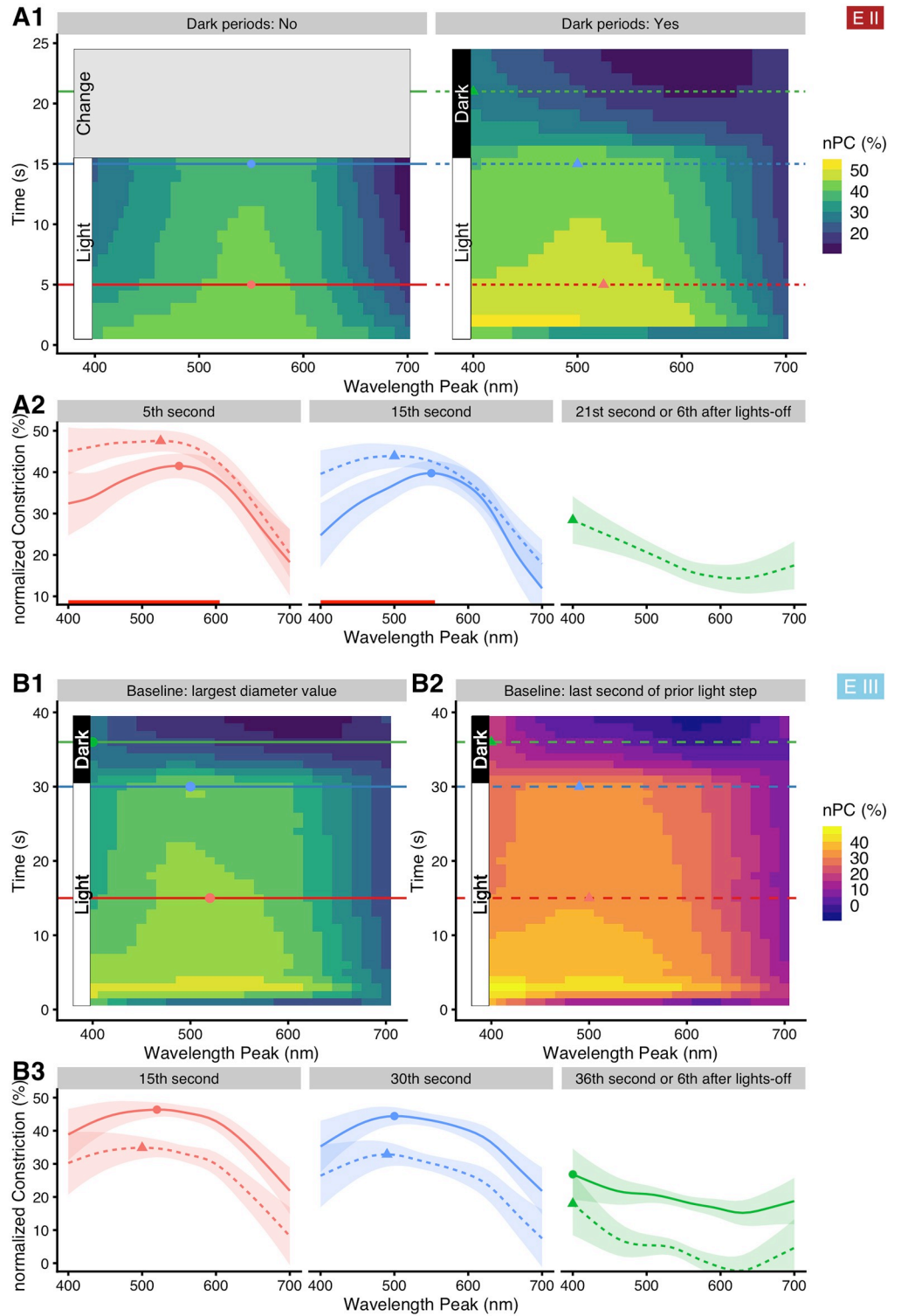
**Other dependencies.** We tested nPC for an interaction of *wavelength* and *series*, which turned out not significant ( $p > 0.18$ ). An influence of the prior wavelength on nPC was not significant either ( $p = 0.37$ ).

Furthermore, we used the average pupil diameter of the last second before the start of a light step as an alternative baseline value in Eq 1. We explored this variant as the response variable in the *wavelength-by-time* model for protocols using periods of darkness between light steps, where this nPC calculation is a valid alternative. As expected, *series* is no longer significant in this case ( $p = 0.31$ ), since it is accounted for by the respective alternative baseline diameter. Irradiance is not significant either ( $p = 0.37$ ). Otherwise, the model does not seem to behave differently in terms of *wavelength* or *time* (see S9 Fig), other than that nPC values are, on the average, about 10% lower.

For the protocols having periods of darkness between light steps, the median standard deviation for the pupil diameter at the last second of darkness prior to each light step was 0.37 mm (range 0.21 to 0.95 mm). This was calculated on a by-participant basis, without model fitting.

**Illuminance.** To analyze how well each type of illuminance can predict nPC in those protocols with periods of darkness between light steps, we constructed simplified standard linear mixed-effect models analogous to those used in *Experiment I*. The results were shown above in





**Fig 8. Interaction of wavelength with time in Experiment II and III.** Time denotes the time (in seconds) since light onset, or light change to the respective wavelength. (A1) and (A2) show results from *Experiment II*, (B1), (B2), and (B3) those from *Experiment III* for the protocols *Long 1* and *Long 2*. (A1) False-color graph of model predictions for the nPC's dependence on wavelength (x-axis) and time (y-axis) for settings with (right panel), or without (left panel), periods of darkness between changes of wavelength. All other predictors are held constant at their average value. Horizontal lines indicate where the respective traces shown in figure part (A2) are taken from; continuous curves refer to the absence of dark periods, dashed curves to their presence; filled circles and triangles mark the wavelength of the



respective maximum nPC value of these traces. In the right panel of A1, nPC values after the 15<sup>th</sup> second (i.e., in the Dark period) show a time-by-wavelength rendition of the post-illumination pupil reflex (PIPR). (A2) Model predictions for nPC vs. wavelength at three points in time. The right panel shows the wavelength dependency of the 6-second PIPR. Ribbons show the 95% confidence interval for the predicted means. Dotted lines with a triangle symbol represent the discontinuous setting with periods of darkness present between light steps; full lines with a filled circle show the continuous setting. The red horizontal line above the x-axis in the two left panels shows where the difference between the two settings is significant at the 0.05 level. (B1) Like (A1 right panel), but for *Experiment III*. Horizontal lines show where the respective traces shown in figure part (B3) are taken from, and points mark the wavelengths of the respective maximum nPC value of these traces. (B2) Like (B1), but with the nPC baseline taken from the last second of the respective previous light step. (B3) Like (A2), but dotted lines and triangles represent the scaling according to (B2), full lines and points according to (B1). The right panel shows the wavelength dependency of the 6-second PIPR.

<https://doi.org/10.1371/journal.pone.0253030.g008>

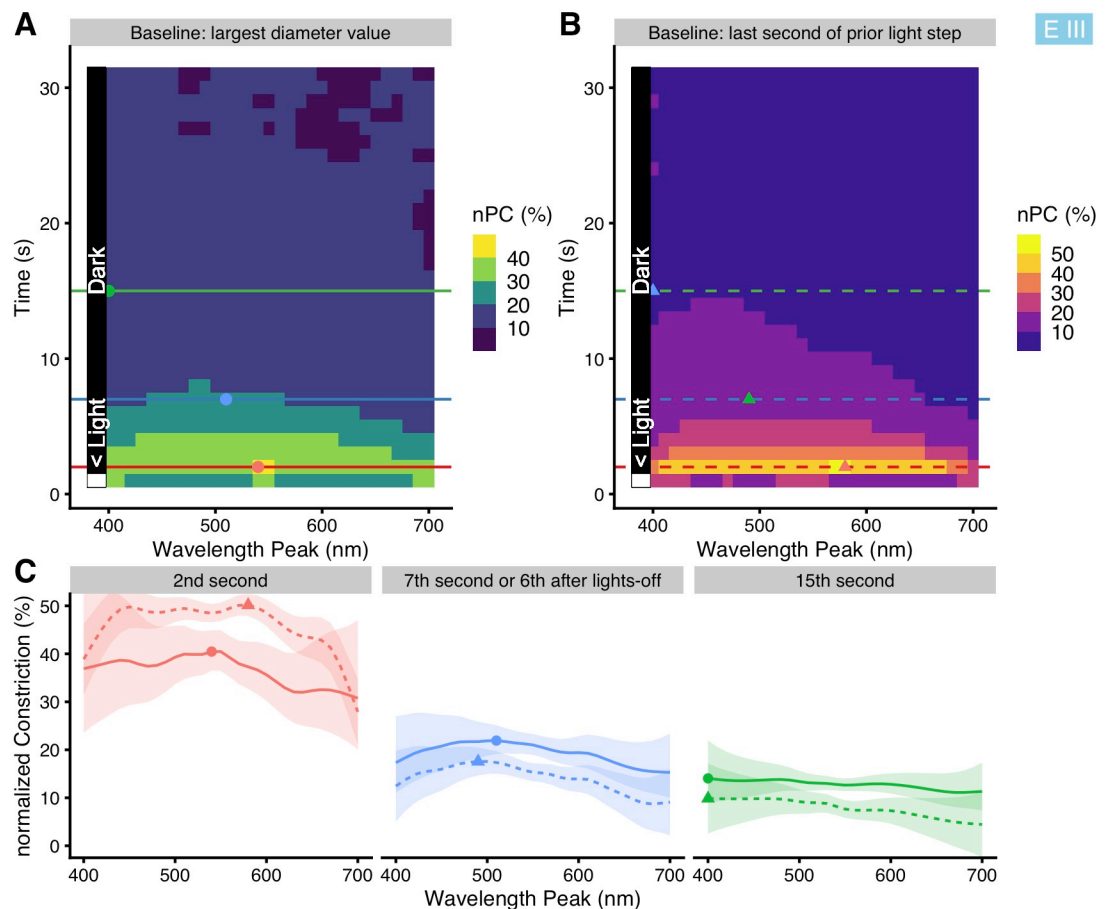
**Fig 7B** (for comparison to *Experiment I*). As in *Experiment I*, they show that mesopic weighing predicts nPC best ( $R^2_{\text{illu}} = 0.55$ ). Photopic functions do somewhat less well in comparison ( $R^2_{\text{illu}} = 0.49$  for  $V_{10}(\lambda)$ , and  $R^2_{\text{illu}} = 0.42$  for  $V(\lambda)$ ). Melanopic and rhodopic receptor functions work better than in *Experiment I* ( $R^2_{\text{illu}} \leq 0.31$ ). Of the three cone functions, chloropic weighing led to the best results ( $R^2_{\text{illu}} = 0.54$ ), with the erythropic function second ( $R^2_{\text{illu}} = 0.43$ ). In contrast, cyanopic weighing was the only illuminance type that had no support by the model at all ( $p = 0.41$ ,  $R^2_{\text{illu}} < 0.01$ ).

## Experiment III

**Base model results.** For the analysis of Experiment III, we analyzed nPC results for the protocols *Short 1* and *Short 2* in models separate from those for *Long 1* and *Long 2*, since the two protocol pairs differ conceptually. For those with one second of light followed by thirty seconds of darkness (*Short 1* and *Short 2*), nPC was calculated as the average nPC value of the sixth second after lights-off (or seventh second of the respective light step). The model shows neither support for an effect of *wavelength*, nor for one of *irradiance* (all  $p > 0.31$ ).

For the two protocols with thirty seconds of light followed by nine seconds of darkness (*Long 1* and *Long 2*), nPC was calculated, as above, as the average nPC value of the last five seconds of light; the results were shown above in [Fig 5C1](#). *Wavelength* has support from the model ( $\Delta\text{AIC} = 22$ ). The predictor shows an inverted-U shaped curve, with a peak at 500 nm and only slightly lower amplitude at short wavelengths ([Fig 5C1](#)). The effect of *series* is not significant ( $p = 0.54$ ) ([Fig 5C2](#)); nPC does not seem to depend on the stimulus' position in the series when the stimulus duration is extended to thirty seconds. In contrast, nPC increases strongly with *irradiance* ( $\beta_{\log_{10}(\text{irradiance})} = +19.7\%$ ,  $\pm 4.3$  SE). *Irradiance* has support ( $\Delta\text{AIC} = 19$ ) as part of the model ([Fig 5C3](#)). Compared to other estimates of *irradiance*, the effect is rather large. See below for a discussion of a possible overcompensation. When removing *irradiance* as a predictor from the model ([S11C Fig](#)), the lower wavelengths become far less influential, with results similar to those in *Experiment I* ([S11A Fig](#)). The standard deviation of the random intercept by participant,  $\sigma_b^2$ , was 5.4% in *Experiment III*.

**Time-course of the wavelength dependency.** Analogous to the analysis in *Experiment II*, we used one-second averaged values in Experiment III to explore how nPC vs. *wavelength* changed over the time course of the light application. The results are shown below in [Fig 9](#) for the protocols *Short 1* and *Short 2*, and were shown above in [Fig 8B](#) for the protocols *Long 1* and *Long 2*, respectively. For the settings with one second of light, followed by thirty seconds of darkness (*Short 1* and *Short 2*), model diagnostics for nPC with the global pupil baseline per protocol were initially problematic, showing heteroscedasticity and skew in the residuals. The skew was eliminated by a logarithmic transformation of nPC and heteroscedasticity was reduced by a logarithmic transformation of *time*. *Time* is the time (in seconds) since light



**Fig 9. Interaction of wavelength with time in Experiment III, for protocols Short 1 and Short 2.** Time is the time (in seconds) since light onset or light change to the respective wavelength. (A) False-color graph of model predictions for the nPC's dependence on wavelength (x-axis) and time (y-axis). All other predictors (basic model) are held constant at their average. Horizontal lines show where the respective traces shown in part (C) are taken from, and points mark the wavelengths of the respective maximum nPC value of these traces. (B) Like (A), but with a different measure of nPC where the nPC baseline is taken as the mean pupil diameter across the last second of the respective previous light step. Model diagnostics are superior to those for the Model in part (A). (C) Model predictions for the nPC vs. wavelength at three points in time after light-step onset: in the 2<sup>nd</sup> second (red traces), the 7<sup>th</sup> second (green traces), and the 15<sup>th</sup> second (blue traces); the latter two cases are also the 6-second PIPR and 14-second PIPR, respectively. Ribbons show the 95% confidence interval for the predicted means. Dotted lines with a triangle represent the scaling according to (B), full lines with a filled circle according to (A).

<https://doi.org/10.1371/journal.pone.0253030.g009>

onset or light change to the respective wavelength. We also analyzed the data using the alternative nPC pupil baseline defined as the average diameter of the last second of the respective previous light step instead of the global baseline diameter taken across all light steps in one protocol. With the alternative baseline, model diagnostics were far more satisfactory compared to the global baseline, and only required logarithmic transformation of time, which is why this model is preferred. Both models used *adaptive splines* for accommodating the rapid changes during the first half of the light step compared to almost no changes during the second half. The results are shown in Fig 9A and 9B, with exemplary sections in Fig 9C. Both models predict that shorter wavelengths lead to a speed decrease in pupil re-dilation compared to longer wavelengths. The sixth-second of the PIPR (Fig 9C, middle panel) shows a peak constriction at 490 nm for the preferred model, and 520 nm for the standard baseline. At the 15<sup>th</sup> second (Fig 9C, right panel), the peak has disappeared. The effect of irradiance is not significant in the preferred model ( $p = 0.88$ ).

For the settings with thirty seconds of light followed by nine seconds of darkness (*Long 1* and *Long 2*), model diagnostics for both baseline methods were satisfactory. The results for these models were shown above in Fig 8B1 and 8B2, respectively, with exemplary sections in Fig 8B3. Both models predict that shorter wavelengths become more important with *time*, compared to longer wavelengths. With the standard baseline (Fig 8B1), peak nPC is at 520 nm after 15 seconds, and at 500 nm at the end of the light application. With the alternative baseline (Fig 8B2), peak nPC is at 500 nm after 15 seconds, and at 490 nm at the end of the light application. Re-dilation after lights-off is similar to that in *Experiment II* (Fig 8A). nPC increases with *irradiance* (Baseline: largest diameter value;  $\beta_{\log_{10}(\text{irradiance})} = +13.0\% \pm 3.1\%$  SE). When leaving *irradiance* out of the model, nPC sensitivity overall shifts slightly towards higher wavelengths (about 10 to 30 nm; S10C and S10D Fig), similar again to *Experiment II*.

**Other dependencies.** We tested for an influence of the prior wavelength on nPC, which was not significant ( $p = 0.68$ ). For the short protocols, the median standard deviation for the pupil diameter at the last second of darkness prior to each light step was 0.48 mm (range 0.25 to 0.76 mm). For the *Long* protocols, it was 0.49 mm (range 0.32 to 1.26 mm). These values were calculated individually for each participant, without model fitting.

**Illuminance.** To analyze how well each type of illuminance can predict nPC in the *Long* protocols (thirty seconds of light followed by nine seconds of darkness), we constructed simplified standard linear mixed-effect models analogous to those in *Experiment I* and *Experiment II*. The results were shown above in Fig 7C but are reported here. As in the other experiments, mesopic weighing predicts nPC best ( $R^2_{\text{illu}} = 0.48$ ). Photopic functions do less well in comparison ( $R^2_{\text{illu}} = 0.45$  for  $V_{10}(\lambda)$ , and  $R^2_{\text{illu}} = 0.40$  for  $V(\lambda)$ ). Curiously, melanopic and rhodopic receptor functions are less good predictors than in *Experiment I* or *Experiment II* ( $R^2_{\text{illu}} \leq 0.21$ ). Of the three cone functions, chloropic weighing led to the best results ( $R^2_{\text{illu}} = 0.44$ ), with the erythropic function second ( $R^2_{\text{illu}} = 0.42$ ). In contrast, cyanopic weighing was the only illuminance type with no support by the model ( $p = 0.24$ ,  $R^2_{\text{illu}} < 0.01$ ).

## Pooled data

To gain further insight into the interaction of *wavelength*, *chronotype*, and *time of day*, we repeated the analysis from *Experiment I* with pooled data from all experiments. Data from protocols *Short 1* and *Short 2* were excluded since these differ conceptually from the others. To partly account for differences between the experiments, we allowed *wavelength* to vary based on experiment type, which is a slight but acceptable oversimplification for *Experiment II* (see above). We further allowed *series* to change, based on the experiment. The interaction effect had strong support ( $\Delta\text{AIC} = 163$ ); results were shown above in Fig 6B. Compared to the analysis based solely on the results of *Experiment I* (Fig 6A), the basic patterns are similar for the evening and the afternoon. Before midday, however, *Larks* do not shift in wavelength sensitivity towards the shorter wavelengths as was the case in *Experiment I*. Further, the shift towards longer wavelengths for *Owls* is not as strong. Lastly, there is strong support of separate interaction effects (*wavelength*, *chronotype*, and *time of day*) based on whether or not there are periods of darkness between light steps ( $\Delta\text{AIC} = 176$ ), but the resulting model predictions do not lead to any insights beyond what is shown in Fig 6. Possible reasons for the differences between the models will be discussed in the next section. As a side note, the effect of *irradiance* was estimated to be an increase of nPC of 8.9% per order of magnitude ( $\beta_{\log_{10}(\text{irradiance})} = +8.9\% \pm 1.3$  SE) irradiance increase.

## Discussion

### Base model results and time-course of the wavelength dependency

Our initial hypothesis was mostly based on findings from McDougal and Gamlin [13] and Gooley et al. [14]. For a continuous series of narrowband light steps, lasting 15 seconds each, we expected cones to contribute minimally after 10 seconds [13]. Furthermore, because of the whole series' considerable length (just over 15 minutes), we expected sluggish ipRGC influences to manifest themselves [14]. Therefore, we believed wavelengths between 490 nm to 510 nm to have a considerable effect on normalized pupillary constriction (nPC) ten to fifteen seconds into each light step. However, the results of *Experiment I* do not support this hypothesis (Fig 5A1). The peak nPC around 540 nm is close to the peak of the long and middle-wave cone functions (L+M), which suggests strong cone influences. There is further support for an interaction of *wavelength* and *series* in *Experiment I* which indicates a slightly heightened nPC at, and shortly after, reaching the wavelength of peak ipRGC sensitivity, but the effect is small to begin with and disappears over the course of the protocol (S2 Fig). Compared to other research [4–6, 10–14], only Alpern and Campbell (5) report similar results (532 nm peak for the 2.5 mm contraction threshold), while most other papers suggest a peak between 480 and 510 nm [4–6, 10–14], depending on experimental conditions.

There are certain differences and limitations when comparing our results with published literature. Firstly, as stated above, we used continuous light in the first protocols, i.e., participants were pre-adapted to light for all but the very first light step. In the literature, most stimuli are presented singly, with some period of dark adaptation in between [19]. This difference in adaptation also influenced the pupil baseline, as discussed above under *Materials and methods*. Secondly, due to our setup pupil reaction was evaluated at some fixed stimulus intensity, while, in contrast, in most publications stimulus intensity is varied to reach a fixed psychophysiological threshold (e.g., 50% constriction) [4]. Thirdly, we operated in the mesopic stimulus range (see above), a limitation of our light source. Although the stimuli around the ipRGC sensitivity maximum can be considered above the ipRGC-sensitivity threshold as shown above under *Materials and methods*, stimuli of about one degree of magnitude higher intensity would have been desirable to elicit stronger ipRGC reactions [21]. Fourthly, differences in *irradiance* between wavelengths required that this variable is included as a predictor in the model, which is relevant in particular for wavelengths at and below 450 nm. Lastly, the long stimulation periods with minor wavelength changes in *Experiment I* made time-dependent confounding influences more likely, e.g., from desensitization or changes in alertness. Therefore, the *series* covariable was included, as the global effect can account for any nonrandom changes across participants over the course of the protocol. Part of the intent of the second and third experiment was thus whether differences in our results from published literature are due to our setup, or due to our protocol, rather than data analysis. Since our initial hypothesis about ipRGC influence on the normalized pupillary constriction (nPC) for a series of narrowband light stimuli did not hold, we aimed to explore several dependencies with the available data and conducted two additional experiments aimed at specific questions arising from our experimental outcomes.

Light-step changes in the first experiment were small, with a wavelength shift of only 5 nm. We set up *Experiment II* to ensure that the wavelength dependency of the first experiment was not some artefact from the near-continuous sweep across the visible spectrum. Results revealed that this was not the case (compare the red trace in Fig 5B1 to the trace in Fig 5A1). More importantly, results showed that short wavelengths had a far stronger influence when the light stimulus was discontinuous between light steps, compared to continuous light (Fig 5B1, blue trace compared to the red trace). While the results for discontinuous light, i.e., with short

periods of darkness between light steps, suggest a rod rather than an ipRGC influence, they agree far better with published findings [4–6, 10–14] than those from the first experiment. Regarding the time shift in wavelength dependency throughout each light step, the curve remains centered at around 550 nm with continuous light, but is shifted towards shorter wavelengths with discontinuous light (Fig 8A2, full vs. dashed curves). In *Experiment III*, thirty seconds of light were followed by nine seconds of darkness. In this case, the last five seconds of light showed even more influence of short wavelengths than in the previous case (Fig 5C1, compared to the cyan trace in Fig 5B1), and the broad peak resembles the behavior shown by Mure et al. [12] for the 0–30 second condition. The shift in wavelength dependency is also more pronounced in *Experiment III* than in *Experiment II* (Fig 8B1 and 8B3, compared to Fig 8A).

In summary for our results, the cone influence fades in favor of rod influence for discontinuous series of long-lasting narrowband light steps, agreeing with findings from McDougal and Gamlin [13]. However, cone influence does not fade for pre-adapted participants as part of a continuous series of light steps, which to our knowledge has not been reported before. Mure et al. [12] showed changes in wavelength dependency occurring with pre-exposure, but this applied for high-intensity stimuli with several minutes of darkness in between. Joyce et al. [60] showed the effect of short-term light adaptation on the PIPR, but not on the light-adapted pupil itself. The continued bleaching of the sensitive rod photoreceptors might desensitize the rod channel, leaving only the changes in cone input as contributors to nPC. While this mechanism cannot be deduced from our data, it seems a plausible explanation and would result in the shown spectral dependency. It was suggested by McDougal and Gamlin [13], however, that the rod channel remains relevant over minutes of continuous light. Instead, we saw that the spectral dependency of the pupillary reaction to narrowband light is similar to the spectral dependency of polychromatic light [16] when narrowband stimuli are applied under pre-adaptation. In vision, interaction effects of rod and cone inputs that depend on the adaptation state of receptors are well known (for a review, see Zele and Cao [61]). In any case, and despite the mesopic experimental conditions, we would have expected some form of ipRGC contribution, which was not apparent, however. Spitschan et al. [62] showed an S-cone opponency to ipRGC input in the pupillary light response, where increasing S-cone stimulation reduced nPC, all other receptor types being stimulated at a constant rate. If cones remain more relevant under pre-adaptation to light compared to prior darkness, as our results suggest, S-cone opponency might help to explain the reduced short-wavelength sensitivity in the presence of relevant ipRGC input to pupil control. IpRGC influence might still be visible in our data, however, as discussed below in *Chronotype and time of day*.

By their presence of a dark period after light stimulation, *Experiment II* and *III* allowed for the analysis of the post-illumination pupil response (PIPR). Surprisingly, the shortest wavelengths led to the PIPR's slowest pupil re-dilation with 15 seconds of prior light in *Experiment II* (Fig 8A, S8 Fig), and even still with thirty seconds of prior light in *Experiment III* (Fig 8B, S10 Fig). The light in the protocols mentioned above was designed to be as continuous as technically possible, leaving only nine seconds of darkness between light steps in the second and third experiment. As can be readily seen in Fig 8, this was not enough time for the pupil to reach a dark-adapted state. Undoubtedly the absence of full dark adaptation influenced the first seconds of the respective following light step. However, we did not find any systematic influence of the prior wavelength on the last five seconds of the respective following light step. Additionally, after *Experiment I*, the sequences of wavelengths were designed to include large and small light-step differences from across the available spectrum. Therefore, any effect that the respective previous light step has is not specific to a particular wavelength but is general to the protocol. This can be seen, e.g., in S9C Fig or Fig 8B3. In these figures, the only difference



is a shift in the y-axis between using an overall baseline or using a baseline from the last second of the previous light step.

Gamlin et al. [10] reported that a Vitamin-A<sub>1</sub> pigment nomogram with peak sensitivity at 482 nm fitted their data closely for the PIPR of two human subjects after only ten seconds of light. However, when looking at the individual data points for *log relative sensitivity* of the human pupil from that study below 482 nm, i.e., at 452 and 473 nm, their sensitivity does not seem to drop far below that of their closest neighboring data point (493 nm). Park et al. [21] showed that the ipRGC influence is best seen during the sixth second of re-dilation after a one-second light stimulus (compared to a ten-second stimulus in that study). The protocols *Short 1* and *Short 2* in *Experiment III* were constructed accordingly (results shown in Fig 9 and S10 Fig). We see a peak nPC at around 490 nm, which would fit the ipRGC sensitivity maximum for a young adult [32]. These results agree with a publication by Adhikari et al. [63]. That study showed an apparent peak sensitivity between 464 and 508 nm (they fitted a Vitamin-A<sub>1</sub> pigment nomogram with peak sensitivity at 482 nm), with an order of magnitude lower relative sensitivity at the 409-nm wavelength stimulus. After the sixth second in our results, however, sensitivity further shifts towards shorter wavelengths. Thus, our results on the PIPR do not contradict published literature but rather suggest a stronger influence of shorter wavelengths than previously thought, below the ipRGC peak at around 490 nm. A possible mechanism might be an S-cone influence on the pupil [64, 65], which has been shown to influence other ipRGC-dependent effects, such as circadian alignment in mice [66] or, very recently, acute Melatonin-suppression in humans [67]. Circumstances for these additional influences include low light levels and comparatively short stimulus times (less than half an hour), which would fit our setup. The S-cone influence on melanopic effects is not undisputed, however, as Spitschan et al. [68] found no evidence for the influence in acute neuroendocrine and alerting responses. A more far-fetched, but possible, mechanism is the influence of other opsin types in the retina, such as neuropsin (Opn5), which have been found in the retina of vertebrates (for a review see Guido et al. [69]).

Regarding the series effect, there seems to be a cumulative effect in *Experiment I* and *II* for the first 18 steps (Fig 5A2 and 5B2), where nPC at first increases with each consecutive light step. The downward slope in *Experiment I* then after about five minutes (Fig 5A2) might be the result of desensitization, where the small differences in wavelength between light steps are not pronounced enough to counter pupillary escape over time. It can be assumed that changes in arousal, which affect the pupil [70], happen across the timespan of the protocols. In Gooley et al. [14], large fluctuations during the ninety-minute light stimulation can be readily seen. In our study, as far as there is some common change in arousal across the participants over time, this effect is integrated into the series effect by the GAMM. As far as there are individual fluctuations in arousal over time, the effect on the pupil will be part of the random smooth per participant. In both cases, the effect of arousal cannot be separated from other effects that contribute to the smooth of *series* or the random smooths. In *Experiment II*, wavelength differences between light steps were larger, and nPC after step 18 did not noticeably increase further, nor decrease as in *Experiment I* (Fig 5B2). Data analysis for *Experiment I* had suggested an interaction of *series* with *wavelength*. However, further analysis had shown that this effect is of little relevance to the interpretation of the data (S2 Fig) and does not suggest ipRGC influence as initially assumed in *Materials and methods*. We also found no indication of influences from the prior wavelength to the current wavelength in *Experiment II*.

The estimated effect of *irradiance* varied between experiments and depended on whether five-second (base model) or one-second averages (time-shift model) of nPC were analyzed. While this is plausible insofar as the experiments differ conceptually, it is still worth discussing the estimates, summarized in Table 1. Except for the *Short* protocols in *Experiment III*,

**Table 1. Model estimates for irradiance across all experiments.**

	Base model	Time-shift model
Experiment I	6.5±1.4%	-
Experiment II	10.7±3.4%	9.0±3.4%
Experiment III	19.7±4.3%	13.0±3.1%
Pooled Data	8.9±1.3%	-

Values denote model estimates of  $\beta_{\log_{10}(\text{irradiance})} \pm \text{SE}$ , which indicates the effect of  $\log_{10}$  transformed irradiance on the normalized pupillary constriction (nPC). The base model uses the average nPC of the last five seconds of light during each light step, the time-shift model uses one-second means of nPC along the time course of each light step.

<https://doi.org/10.1371/journal.pone.0253030.t001>

*irradiance* was a significant predictor in all cases. In the *Short* protocols, the *irradiance* differences between the light steps were seemingly not big enough to influence re-dilation significantly. For the other cases, the estimated value is smallest for *Experiment I*, which is plausible since the changes in *irradiance* between consecutive light steps are the smallest across the experiments. In *Experiment II*, the effect of *irradiance* is similar between the base model and the time-shift model. In *Experiment III*, the estimated effect of *irradiance* for the base model is rather large compared to the other experiments and the time-shift model. The estimate for the time-shift model in *Experiment III* is closer to estimates in *Experiment II*. We believe the estimate for the base model of *Experiment III* to be unlikely high and see no theoretical basis for the large difference to the other estimates. This would mean that the *wavelength* effect of the base model in *Experiment III* overcompensates for *irradiance*. Therefore, the time-shift model is to be preferred to the base model, but the interpretation of *wavelength* as discussed above does not change in a relevant manner (compare Fig 8B3: 30<sup>th</sup> second to Fig 5C1). For the model from pooled data of all experiments, estimates for *irradiance* were close to *Experiment II*. Setting the base model in *Experiment III* aside in favor of the time-shift model (see above), the differences between and within experiments stay broadly the same regardless of the estimated value of *irradiance*, thereby retaining the interpretations above regarding nPC's dependency on *wavelength*.

Finally, regarding the base models, the nPC-by-participant random effect in all three experiments show the large interindividual differences in pupil reaction, even after normalization. These interindividual differences are often remarked in other publications, as summarized by Kelbsch et al. [19]. Depending on the experiment, the average nPC is predicted by the model to vary by about 20% between two extreme participant cases (0.025 to 0.975 percentile difference).

## Chronotype and time of day

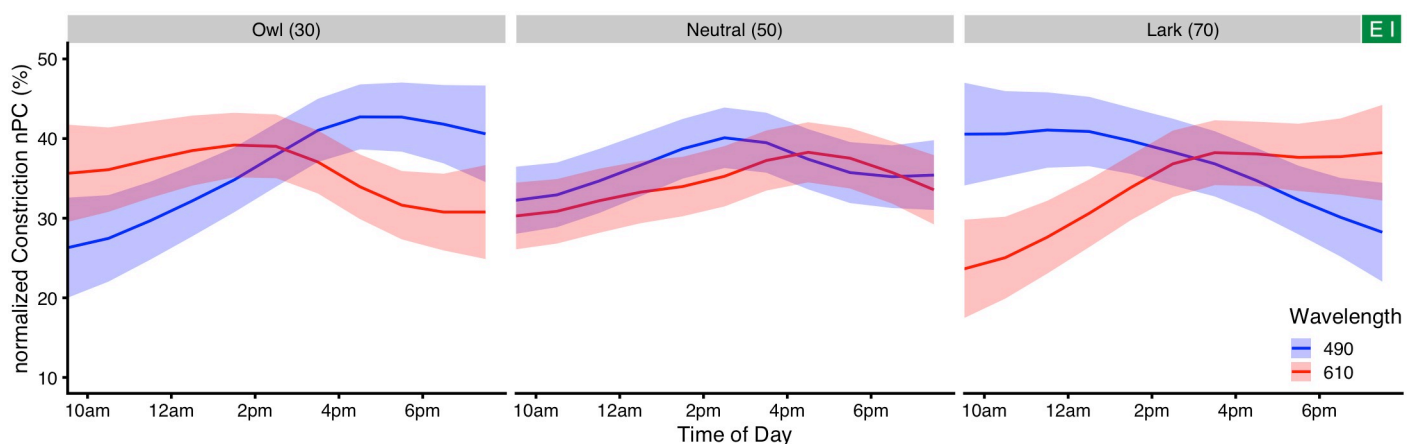
The models from *Experiment I* and pooled data across all experiments strongly suggest the presence of an interaction effect that influences nPC's *wavelength* dependency by an individual's *chronotype* and the measurement's *time of day*. Specifically, early chronotypes, or *Larks*, before noon seem to have a heightened sensitivity to shorter wavelengths (Fig 6A1 and 6B1). In contrast, late chronotypes, or *Owls*, before noon seem to have a slightly lowered sensitivity to shorter wavelengths. These chronotype-specific differences in sensitivity before noon are called into question by the model constructed from pooled data across the experiments (Fig 6B1 and 6B2). Both models agree, however, that in the early evening (6:30), the wavelength of maximum sensitivity shifts to longer wavelengths for the *Larks*. For them, across the day the maximum sensitivity shifts from the longer to the shorter wavelengths, opposite to the

behavior in *Larks*. In general, the shifts are predicted to be stronger for the more pronounced or extreme chronotypes, compared to moderate types. For *Neutral* types, the shift is predicted to be minimal. We also see the chronotype bias in *Experiment I* in the daytime scatterplot (Fig 4B, the green dots show *Experiment I*), mostly for *Owls* during the first half of the day, which is very likely due to *Owls*' later preferences. We pooled the data across all experiments in the hope to gain more insight from the available data. The pooling adds fifty protocol-runs to the eighty runs of *Experiment I* and leads to a more well-rounded distribution of chronotypes across the day (Fig 4B, spread of the dots of all three colors). The pooling also introduces two new dependencies that might influence the interaction and need to be considered. Firstly, *Experiments II* and *III* were conducted decidedly later in the year, compared to *Experiment I* (Fig 4C). The thereby introduced general time-shift by the shortened photoperiod should not be too influential according to a recent paper by McHill et al. [71]. In that study, the time shift in dim-light melatonin onset was less than half an hour across three months from February to May. However, the shortened photoperiod might lead to changes in the possible mechanism underlying the interaction effect ([72], see below). Secondly, whether the light stimulus in the experiments is continuous or discontinuous makes a difference in wavelength dependency (Fig 5B1), and likely also with respect to the involved receptor types (see above for the discussion of the base model). Different receptor inputs likely lead to differences in the interaction, which is supported by a model differentiating the interaction of *wavelength*, *chronotype*, and *time of day* based on whether or not there are periods of darkness between light steps. We slightly prefer the simpler model constructed from *Experiment I*, because while pronounced chronotypes before noon are lacking (Fig 4B, green points), it avoids the likely confounding influences from the pooled data mentioned above. It ultimately makes little difference which of the models is preferred, as both indicate the presence of a similar interaction of *chronotype* with *time of day* on the wavelength dependency. The precise nature of this dependency is beyond this study's scope. We checked for possible confounding variables between *chronotype*, *time of day*, and *age*, all of which did not correlate with one another in any relevant manner. There is also the possibility that we introduced confounding elements into the study design. One could be *alertness*, which is *time-of-day* dependent [73], and, as has already been stated, affects the pupil. However, this dependency on its own would just mean that the *time-of-day* effect, as shown in Fig 5D3, were influenced by alertness. Alertness might further influence the interaction of *time of day* with *chronotype*. Yet for alertness to also change the dependency on wavelength would mean changing the underlying retinal neurocircuitry. We are not aware that this is possible. The second confounding option is the effect of prior light exposure in terms of the participants' history of daily light exposure. As the participants arrived in the laboratory approximately thirty minutes before the respective protocol started, they will have *consumed* [74] varying amounts of light on that day. This amount is *time-of-day* dependent [74] and it is reasonable to assume that it is also depends on chronotype. It is further known that the amount of prior light affects non-visual effects of light, such as melatonin suppression [75], and might also influence the retinal neurocircuitry through the dopaminergic link described further below. In summary, the chronotype effect will have to be approached in a controlled experiment, with measurements for all chronotypes evenly distributed across all times of day while controlling for alertness and prior history of exposure to light, among others.

Circadian effects on the pupil response were reported before and we will discuss two relevant publications in the context of our results. While the focus in those studies is different from ours, we can still compare how their results on circadian ipRGC input line up with our circadian effect. This is under the assumption, that our circadian effect is moderated by ipRGC input, as discussed below. It is further important that, in those studies, the effects of chronotype are adjusted for by looking at participants' internal time or circadian phase, instead of at

external time. While this avoids confounding from differences in internal time between participants, it also makes it hard to determine how much of a difference the chronotype signifies. The MEQ, which we used to assess chronotype, does not describe the participant's phase of entrainment of, but rather their time of preference [76]. And while there are correlations between the phase of entrainment and the MEQ score, an exact phase shift cannot be deduced from the score. Therefore, assumed time-shifts between chronotypes in the following section are to be taken tentatively.

Zelevansky et al. [27] showed a circadian response of ipRGC input to the PIPR on eleven participants in a 20–24h laboratory experiment, independent of external light cues. They also investigated circadian changes in the pupil reflex during their ten second light stimulus, but found only growing effects from fatigue owing to the long experiment. IpRGC input was evaluated in that study through the reduction of normalized pupil diameter after lights off (PIPR), compared between a blue and a red narrowband light stimulus (peak wavelength at 488 nm and 610 nm, respectively). The normalized pupil diameter started to decrease about five hours before the evening melatonin onset (DLMO) for the red stimulus, and three hours prior to the DLMO for the blue stimulus, thereby implying a shift in wavelength dependency. The normalized pupil diameter reached its minimum shortly after the DLMO for both wavelengths; the overall period with a reduced pupil diameter lasted longer for the red than the blue stimulus. As mentioned above, any effect on the circadian ipRGC input due to chronotype is not discernible from these results, since the study centered participants' rhythms on their respective points of melatonin onset, thereby eliminating chronotype differences. However, we would expect the effect of chronotype on melatonin onset to be a pure time-shift of a few hours [77]. Translating these findings to our context, we would expect a fixed relationship between nPC curves of comparable wavelengths across the chronotypes, with a general time-shift by about 1.5 to 2 hours between *Owls* and *Neutral* types, and another shift between *Neutral* types and *Larks*. As shown in Fig 10, our analysis rather suggests a crossed response of nPC depending on wavelength and chronotype (cut-off at 8pm, after which we did not collect data). We thus do not believe that the circadian effect described by Zelevansky et al. [27] is the same as that in our data. Since that study focused on the afternoon and evening, the comparison to our data does not change depending on whether the data from *Experiment I* or the pooled data are used.



**Fig 10. Select changes of nPC across the day in *Experiment I*.** Model predictions for the nPC vs. time of day for three chronotypes (left to right), and two wavelengths (blue trace 490 nm and red trace 610 nm), that were chosen for comparison with the studies by Zelevansky et al. [27] and Munch et al. [26]. Ribbons show the 95% confidence interval for the predicted means.

<https://doi.org/10.1371/journal.pone.0253030.g010>

Another study, by Munch et al. [26], also reported circadian effects on the PIPR. For ten subjects throughout two 12-hour sessions, the 6-second PIPR was assessed after 1-second and after 30-second light stimuli. The authors adjusted the circadian phase in a similar manner to Zele et al. [27]. The normalized 6-second post-stimulus pupil size showed a circadian behavior in the short-wavelength, but not the long-wavelength, stimulus. Translating circadian phase and normalized pupil size to our context, nPC for blue stimuli would be highest before noon and decrease afterwards, with *Owls* having a later shift compared to *Larks*. Again, these predictions do not fit our data.

In both of the abovementioned studies, we expect the effect of chronotype to be an effect of time-shift of a few hours [77]. However, our model suggests something different, namely an interaction similar to the so-called synchrony effects of chronotype reported in the context of cognitive performance in education [78, 79], or of physical performance in athletes [80]. Goldstein et al. [79] define chronotype synchrony as the state in which the time of optimal performance is equal to the time of preference of the chronotype, e.g. morning for *Larks* and evening for *Owls*. Our model from *Experiment I* predicts such a heightened sensitivity for short wavelengths at the time of preference for the chronotypes. This connection might even go one step further. It is known that the time-of-day is highly relevant for the magnitude and sign of a circadian phase shift for a given stimulus [81] (for a review of known influences, see Prayag et al. [82]). Furthermore, Roenneberg and Merrow [83] argue that chronotype is not only the result of non-24-hour internal periods but also of the individual's photosensitivity. Following this reasoning, heightened sensitivity to shorter wavelengths at circadian intervals could be part of a reciprocal system to strengthen or even cause a chronotype, rather than being simply a correlation. *Reciprocal* meaning that some mechanism—like an internal period of other than 24h duration—might predetermine a circadian type. Heightened sensitivity to wavelengths for a circadian shift at times of preference (e.g., towards the evening in *Owls*) means one would need more light outside of the preferred time to shift towards a different chronotype. Conversely it would need less light for strengthening the rhythm at the time of preference, thereby solidifying the chronotype. While causation cannot be determined from our study, intraocular mechanisms have been described to allow for such an effect. In general, rod and cone sensitivity follow a circadian rhythm [84, 85], with melatonin and dopamine as key actuators. Furthermore, ipRGCs are intricately and reciprocally connected to dopaminergic amacrine cells (for a review in the context of myopia, where dopamine plays an essential role, see Stone et al. [86]). Through this connection, ipRGCs can attenuate the outer retinal light adaptation in mice [87]. This system is involved in seasonal and circadian regulation of photosensitivity [72]. Other seasonal changes, like that of human color perception [88], might also be connected to this mechanism. Since ipRGC sensitivity is the basis of the mechanism, the shift towards, or away from, the short wavelength spectrum is plausible. Whether the mechanism operates in a chronotype-dependent manner, as shown above, is open for evaluation.

## Sex and age

We saw strong support for the model which included *sex* as a predictor. There were no differences in dependence on *wavelength* related to *sex*. However, women were estimated by the model to have a lower average nPC. Chellappa et al. [25], in contrast, found sex differences in light sensitivity. In their study, men showed a stronger response to blue-enriched light in the sleep EEG (NREM sleep slow-wave activity) and for vigilant attention, and had higher brightness perception during blue-enriched light. Men also found blue-enriched light preferable to non-blue-enriched light, contrary to women. The stronger average nPC for men in our study would align with a higher brightness perception for men per se, but unlike in Chellappa



et al.'s study was not limited to short wavelengths. Since men—in general—are more likely to be late chronotypes compared to woman [89], and Chellappa et al. [25] did not mention whether they stratified their sample for chronotype, their reported effect might also have been caused, in part, by a chronotype effect. That study was performed in the late evening, where our interaction effect of chronotype with time of day and wavelength suggests a heightened sensitivity for short wavelengths (see above).

Age was estimated to influence nPC, but there was no indication of an interaction with *wavelength*. The latter agrees with a publication by Rukmini et al. [90], where pupillary responses to short-wavelength compared to long-wavelength light were independent of age (in a comparison of two age groups: 21–30, and  $\geq 50$ ). Depending on whether or not chronotype and time of day were included in our model, nPC was predicted to rise steadily with age (Fig 5D1, S3A Fig), or reach a plateau at about age 25 (S3B Fig). According to Rukmini et al. [90], pupillary responses were reduced in older participants. Age was also part of the unified model from Watson and Yellott [3] based on data from Winn et al. [91]. Winn et al. [91] focused on pupil diameter itself, rather than the (relative) change in diameter, i.e., nPC. The authors calculated the age-dependent pupil-diameter change per log unit stimulus luminance from five illuminance levels (9 to 4400 cd/m<sup>2</sup>). The resulting regression showed that pupil constriction decreases with age in the photopic range. Conversely, Daneault et al. [36] found no significant age-related differences in pupil constriction, even though pupil size itself decreased with age. Adhikari et al. [92] showed that the relative peak constriction did not change with age, that measure being closely related to our dependent variable nPC. While not significantly correlated with *age*, the trends in their study for both blue and red stimuli show a slight positive slope. In summary, literature reports show a heterogeneous picture of the effect of age on nPC. Our age range was small by design and focused on participants below the age of forty. Accordingly, the effects of *age* in our study were small and appear compatible with any of the studies mentioned above, thereby contributing little for this specific aspect.

### Effect sizes

To summarize the unstandardized effect sizes of the various dependencies (Fig 5), wavelength clearly shows the biggest effect. nPC changes by about 20% across the available spectrum in cases of continuous light (Fig 5A1 and 5B1) and by more than 30% in cases of discontinuous light (Fig 5B1 and 5C1). The long series of small wavelength changes in *Experiment I* also had a considerable effect, of about 13% (Fig 5A2), compared to about 5% for the larger changes in *Experiment II* or no effect for the thirty-second stimulus in *Experiment III*. We already discussed the effect of *irradiance* above. Depending on the model estimate, a 0.5 log<sub>10</sub> difference in *irradiance* changes nPC by about 5% (Fig 5A3, 5B3 and 5C3), which is roughly the same influence that *sex* or *age* had in our setup (see Fig 5D1 for *age*). The estimated main effects of chronotype and time of day amount to about 10% across their respective range. The combined effect of these two covariates and wavelength leads to considerable changes of about 20% nPC (Fig 6A2). As this effect has not been reported before, it may be a consequence of the used protocols for continuous light since the effect is far diminished when evaluated across protocols (Fig 6B2). This is under the assumption that the effect can be replicated without confounding influences as discussed above.

### Illuminance

At present, there are nine distinct measures of illuminance in common use to describe visual and alpha-opic (i.e., single receptor-dependent) effects of light. Scotopic illuminance and rhodopic equivalent daylight illuminance count as two, but both use the same rhodopic action

spectrum. Many publications use  $V(\lambda)$ -derived luminance [3] (in  $\text{cd}/\text{m}^2$ ), but since the spatial stimulus characteristics are unchanged across our experiments (Ganzfeld conditions), luminance and illuminance are interchangeable in terms of comparisons between the types of illuminance. Adrian [7] argued that mesopic weighing was sufficient to explain apparent wavelength-dependency shifts towards the shorter wavelengths (under mesopic conditions). Accordingly, Adrian explained the mechanism behind this shift as purely S-cone and rod-based. We now know that Adrian [7] was incorrect in his assumptions about the sole mechanism and that there is additional input from ipRGCs in the pupillary system, e.g. from the work of Gamlin et al. [10] or Zaidi et al. [11]. However, the appeal of the mesopic weighting compared to others is easy to see across all three of our experiments (Fig 7), where the mesopic illuminance best predicted nPC, thereby considering all cone and rod input. The single-opsin weighting functions for rhodopsin or melanopsin perform far less well in all cases.

In summary, our findings are relevant in the context of arguments made by Spitschan [15] and Zandi et al. [16] about the limited applicability of luminance-based pupil models [3] in the context of ipRGC inputs. At least for our experimental setup of narrowband and mesopic light stimuli of changing wavelength, mesopic weighting was best at predicting pupil constriction. Conventional photopic (il)luminance-based approaches performed still well and, compared to single-opsin based functions, were preferable to predict pupil constriction. Finally, Spitschan et al. [62] reported an S-cone opponency, where the pupil would curiously increase in diameter with increasing stimulation of S-cones. Since our setup did not account for the constant stimulation of all other receptor types as that study did, any opponency effect is expected to be decreased in our study. The cyanopically weighted equivalent of daylight illuminance (S-cone weighing) was not significantly correlated with nPC in any experiment here. The cause is not known, but the results are compatible to a zero-sum of opposing inputs.

## Conclusion

The spectral dependency of pupillary reactions to narrowband light is well understood for isolated stimuli, but less so for scenarios with preadaptation to light. We looked at the normalized pupillary constriction (nPC) in protocols of narrowband light of periodically changing wavelength and found very little influence of short wavelengths. A second experiment showed that this effect originated in the continuous application of light and that the effect was present over the whole time-course of each wavelength. Our results for continuous narrowband light and literature findings from polychromatic light share a similar wavelength dependency, compared to singularly presented narrowband light. Furthermore, the 6-second post-illumination pupil response (PIPR) to isolated light stimuli, for two out of three durations showed that the slowest re-dilation happens for the shortest wavelengths, well below peak ipRGC sensitivity, implying the presence of different, or at least additional, receptor influences to the PIPR. Pupillary reactions to isolated light stimuli and the PIPR to short light stimuli behaved very much as would be expected from published literature. This strengthens our belief that our results are valid despite some technical limitations as described above. We further showed that mesopic illuminance was the best measure of illuminance to explain the pupil reaction. We found dependencies of *sex* and *age* on the pupillary response through the exploration of several covariates, but mainly an interaction of *chronotype* and *time of day* with *wavelength*. The interaction effect implies a modulation of wavelength-sensitivity with a heightened sensitivity to shorter wavelengths at the time of chronotype preference. This circadian effect further seems to manifest itself differently, depending on whether light is continuous or discontinuous. The effect could be linked to a mechanism that strengthens an individual's chronotype at the current time of preference, or it could result from chronotype and time-of-day dependent

differences in the light history. If the effects are reproducible in a controlled experiment, they might even act as a form of marker for the individual's chronotype.

## Supporting information

**S1 Fig. nPC depending on wavelength for different dark adaptation periods (DAPs).** Traces show the mean, LOESS-smoothed nPC vs. wavelength for two subjects, each repeating all shown protocols three times. DAP varied between 1.5, 3, and 15 minutes (lilac, blue, dark green). For comparison, results of a second, consecutive run (also performed three times, light green) of the first protocol is shown, as is the *Down* protocol (yellow). Points show the nPC values from which the traces were constructed. More information is found in *Materials and methods*.

(TIF)

**S2 Fig. nPC depending on wavelength with and without interaction in Experiment I.** Model prediction for models with (dashed lines), and without an interaction effect (solid lines) of *wavelength* and *series*, across the series. The number above each plot indicates the series number. Colored dots represent raw nPC values at the respective series number and wavelength, their color indicates the respective protocol. The positions of points on the x-axis indicate where changes between the two models (lines) should be evaluated. For the model with the interaction effect, it seems as though nPC is increased after reaching the ipRGC peak sensitivity (at about 490 nm for a young adult), but not strongly and not for long.

(TIF)

**S3 Fig. Influence of age in Experiment I.** (A) Model predictions how nPC depends on *age*, when all other predictors (basic model) are held constant at an average level. (B) Like (A), but for the model with all other dependencies in *Experiment I* included. Here, *age* seems to affect nPC with a more continuous rise over the age range.

(TIF)

**S4 Fig. GIF Animation of nPC depending on wavelength, time of day, and chronotype in Experiment I.** Model predictions show time-of-day values at half-hour points from 8:30 am to 7:30 pm. Left: False-color model predictions for hourly nPC values depending on *wavelength* (x-axis) and *chronotype* (y-axis), when all other predictors (basic model) are held constant at an average level. Horizontal lines show where the traces from the right image are taken from. Right: Model predictions for hourly nPC values depending on wavelength for three chronotypes: *Larks* (green traces, CT score 70), *Owls* (red traces, CT score 30), and *Neutral* types (blue traces, CT score 50). Ribbons show the 95% confidence interval for the predicted mean values.

(GIF)

**S5 Fig. nPC depending on wavelength, time of day, and chronotype in Experiment I.** False-color contour-line graphs of nPC model predictions for several wavelengths between 430 and 670 nm (color scale shown in the inset at the upper right corner of each plot). All plots are scaled equally. Each plot visualizes nPC depending on *time of day* (clock time in 24h values, x-axis) and *chronotype* (higher values are morning types, lower values evening types). Plots were created with the *vis.gam()* function in R, with the *too.far* argument set to 0.1. The *too.far* argument excludes grid points from the plot, when points are not represented by variable combinations close enough to actual data. Thereby, *too.far* is a measure of accepted extrapolation, scaled from 0 to 1.

(TIF)

**S6 Fig. Basic model predictions for unscaled changes in pupil diameter, and raw pupil diameter, in *Experiment I*.** Model predictions for response variables other than nPC, versus three main predictors (*wavelength*, *series*, and *irradiance*), when all other predictors are held constant at their average level. Traces show the model prediction for the mean, ribbons its 95% confidence interval. Red dotted lines show particular predictor values for the plotted relationship. (A / D) Response vs. stimulus peak wavelength. (B / E) Response vs. series number of light steps. (C) Response vs. stimulus irradiance. The x-axis scaling reflects the logarithmic transformation of irradiance.

(TIF)

**S7 Fig. GIF Animation of normalized pupillary constriction (nPC) depending on wavelength and time in *Experiment II*.** Left: False-color model predictions for nPC depending on *wavelength* (x-axis) and *time* (y-axis), for settings with, or without, periods of darkness between wavelength steps. All other predictors are held constant at an average level. Horizontal lines show where the traces from the right image are taken from. Right: Model predictions for the nPC vs. *wavelength* over *time*. Ribbons show the 95% confidence interval for the predicted means. Blue lines represent the setting when periods of darkness are present between light steps, red lines when not. The red horizontal line above the x-axis shows where the difference between the two settings is significant (5% level).

(GIF)

**S8 Fig. Interaction of wavelength with time in *Experiment II*.** (A) False-color graph of model predictions for the nPC's dependence on *wavelength* (x-axis) and *time* (y-axis) for settings with, or without, periods of darkness between wavelength steps. All other predictors (basic model) are held constant at their average. Red dots show the peak nPC value for each second, the value right next to it the respective peak wavelength. Green dots and values show the respective trough for nPC. The trough is at 700 nm where green dots are not shown. (B) Like (A), but for the model without *irradiance* as predictor.

(TIF)

**S9 Fig. Interaction of wavelength with time in *Experiment II*, for protocols with darkness between light steps.** (A) False-color graph of model predictions for the nPC's dependence on *wavelength* (x-axis) and *time* (y-axis). All other predictors (basic model) are held constant at their average. Horizontal lines show where the respective traces shown in part (C) are taken from. (B) Same as in (A) except that the baseline for nPC calculation (Eq 1) is taken from the last second of the respective previous light step. (B) The graph serves to check whether and how the single baseline value per protocol (shown in A) changes the model prediction. (C) Model predictions for the nPC vs. wavelength at three points in time after light step onset: in the 5<sup>th</sup> second (red), 15<sup>th</sup> second (blue), and 21<sup>st</sup> second (green); the latter case is also the sixth seconds after lights-off. Ribbons show the 95% confidence interval for the predicted means. Full lines represent nPC values when baseline pupil values are taken from each protocol's largest pupil diameter, dotted lines when baseline pupil values are taken from the last second of the respective prior wavelength-step.

(TIF)

**S10 Fig. Interaction of wavelength with time in *Experiment III*, for models with, or without, irradiance as predictor.** (A) False-color graph of model predictions for the nPC's dependence on *wavelength* (x-axis) and *time* (y-axis) for the protocols with one second of light, followed by thirty seconds of darkness, and with *irradiance* as the predictor. All other predictors (basic model) are held constant at their average. Red dots show the peak nPC value for each second, the value right next to it the respective peak wavelength. Green dots and values show the

respective trough for nPC. The trough is at 700 nm where green dots are not shown (e.g. between 2 and 30 seconds). The baseline for nPC calculation (Eq 1) is taken from the last second of the respective previous light step, since this baseline led to a preferable model in terms of model diagnostics. (B) Like (A), but for the model without *irradiance* as the predictor. (C) Like (A), but for protocols with thirty seconds of light, followed by nine seconds of darkness with irradiance as a predictor. The baseline for nPC is as described in *Materials and methods*. (D) Like (C), but for the model without *irradiance* as the predictor.

(TIF)

**S11 Fig. Model predictions for models without irradiance as a predictor.** Normalized pupillary constriction (nPC) as depending on *wavelength*, when all other predictors are held at an average, constant level. Traces show the model prediction for the mean, ribbons its 95% confidence interval. Dotted lines show the respective peak. (A) through (C) show dependencies in *Experiment I, II, and III*, respectively. They can be compared to the model results which include irradiance in Fig 5A1, 5B1 and 5C1.

(TIF)

**S12 Fig. Experimental data.** (A) Pupil diameter plotted against wavelength, for each of the nine protocols. The color scale shows at which point in the series a specific wavelength was presented; light yellow represents early in the series. For all but the *Short* protocols, points represent the average nPC during the respective last five seconds of a light step. In the two *Short* protocols, points represent the average nPC during the sixth second after lights-off (or seventh second after lights-on). Traces show the mean nPC, ribbons its standard deviation. The number in the upper right corner of each plot shows the corresponding sample size.

(TIF)

**S13 Fig. Plot of the 61 spectral distributions shown in Fig 1A.** Corneal irradiance ( $\text{W}/\text{m}^2$ ) vs. wavelength, labelled by wavelength peaks (404–695 nm) above each subgraph. The x-axis on the bottommost plots and the y-axis on the left apply to all plots in their respective row and column. Blue and red traces show the spectral measurements with and without an estimate of prereceptor filtering, respectively. Wavelength peaks above each plot state the peak for the raw stimulus measurement. All displayed values are based on spectral irradiance measurements with a field-of-view restriction according to the CIE S 026 standard [32]. In our case, these measurements are 24% lower than those of the unobstructed sensor diffusor.

(TIF)

**S14 Fig. Mean irradiance and standard deviation.** Corneal irradiance ( $\text{W}/\text{m}^2$ ) against wavelength for all wavelength peaks, with their respective standard deviation across multiple measurements. All displayed values are based on spectral irradiance measurements with a field-of-view restriction according to the CIE S 026 standard [32]. In our case, these measurements are 24% lower than those of the unobstructed sensor diffusor.

(TIF)

**S1 File. R-Markdown script results for calculating nPC from measurement data.** Zip file containing nine html files with results from R-Markdown scripts. The scripts are one each for the nine protocols used across the experiments. Since the raw measurement data are not part of the provided supplements, we included one exemplary result file for each protocol instead of the script itself. These html files show the code, as well as the output for the example participant.

(ZIP)



**S2 File. Irradiance measurements and illuminance values for the range of light stimuli.**

Microsoft Excel file containing two worksheets. The worksheet 'Measurements' contains the irradiance measurements in a resolution of 1 nm. The worksheet 'Illuminance' contains all types of illuminance values described in the main text and irradiance. All displayed values are based on spectral irradiance measurements with a field-of-view restriction according to the CIE S 026 standard [32]. In our case, these measurements are 24% lower than those of the unobstructed sensor diffusor.

(XLSX)

**S3 File. R-Markdown script used for statistical analysis and graphics generation.** Zip file containing several files necessary to replicate the statistical analysis and generation of graphics. Besides some external tables, an R-function file, and three pictures to mark the respective experiment in graphs, the zip file consists of nine R-Markdown scripts. One of these is used to set up the data prior to analysis. The necessary data can be downloaded from the *Open Science Framework* [39]. Five of the R-Markdown files are for analysis of *Experiment I, II, III Short, III Long*, and the pooled data. The eighth file is for graphics generation. The second-to-last file takes a sample of ten random participants from the first experiment for every protocol and builds the base model from this subset. This sample analysis is to show that the main dependencies of *wavelength* and *series* can come from a smaller sample. An html file with the same file name shows an example. The last file is for calculating estimates for prereceptor filtering and using those estimates to calculate irradiance and photon density values from the spectral irradiance measurements.

(ZIP)

**S4 File. Html file with 3-D visualizations of normalized pupillary constriction (nPC)**

**depending on wavelength and time in Experiment II.** The zip file contains three files. One html file with a scalable and rotatable diagram of nPC (denoted as "Amplitude") vs. wavelength and *time*. An html viewer, or browser with support for html widgets, is needed. The other two files are snapshots from the html file for each condition of *Dark*. Color mapping is according to the z-axis (nPC) for better visibility.

(HTML)

## Acknowledgments

We thank Jorge Alberto Gutiérrez Alvarez, Theresa Scherzer, Daniel Philipp Setzensack, and Regina Heiß for their help with data collection as part of their respective thesis work. We thank Moritz Faust for his support on data collection and questionnaire digitalisation. We also thank three anonymous reviewers for their helpful and constructive input to the manuscript.

## Author Contributions

**Conceptualization:** Johannes Zauner.

**Data curation:** Johannes Zauner.

**Formal analysis:** Johannes Zauner.

**Funding acquisition:** Johannes Zauner.

**Investigation:** Johannes Zauner.

**Methodology:** Johannes Zauner.

**Project administration:** Johannes Zauner.

**Resources:** Herbert Plischke.

**Supervision:** Herbert Plischke, Hans Strasburger.

**Validation:** Johannes Zauner, Hans Strasburger.

**Visualization:** Johannes Zauner.

**Writing – original draft:** Johannes Zauner.

**Writing – review & editing:** Johannes Zauner, Herbert Plischke, Hans Strasburger.

## References

1. McDougal DH, Gamlin PD. Autonomic Control of the Eye. *Comprehensive Physiology* 2014. p. 439–73.
2. Szabadi E. Functional Organization of the Sympathetic Pathways Controlling the Pupil: Light-Inhibited and Light-Stimulated Pathways. *Front Neurol.* 2018; 9:1069. Epub 2019/01/09. <https://doi.org/10.3389/fneur.2018.01069> PMID: 30619035.
3. Watson AB, Yellott JI. A unified formula for light-adapted pupil size. *J Vis.* 2012; 12(10):12. Epub 2012/09/27. PMID: 23012448.
4. Wagman IH, Gullberg JE. The relationship between monochromatic light and pupil diameter. The low intensity visibility curve as measured by pupillary measurements. *Am J Physiol.* 1942; 137(4):769–78. <https://doi.org/10.1152/ajplegacy.1942.137.4.769>
5. Alpern M, Campbell FW. The spectral sensitivity of the consensual light reflex. *J Physiol.* 1962; 164(3):478–507. <https://doi.org/10.1113/jphysiol.1962.sp007033> PMID: 14012269.
6. Bouma H. Size of the static pupil as a function of wave-length and luminosity of the light incident on the human eye. *Nat.* 1962; 193(4816):690–1. <https://doi.org/10.1038/193690a0> PMID: 13871842
7. Adrian W. Spectral sensitivity of the pupillary system. *Clin Exp Optom.* 2003; 86(4):235–8. <https://doi.org/10.1111/j.1444-0938.2003.tb03111.x> PMID: 12859242
8. Berson DM, Dunn FA, Takao M. Phototransduction by retinal ganglion cells that set the circadian clock. *Sci.* 2002; 295(5557):1070–3. <https://doi.org/10.1126/science.1067262> PMID: 11834835
9. Lucas RJ, Douglas RH, Foster RG. Characterization of an ocular photopigment capable of driving pupillary constriction in mice. *Nat Neurosci.* 2001; 4(6):621–6. <https://doi.org/10.1038/88443> PMID: 11369943
10. Gamlin PD, McDougal DH, Pokorny J, Smith VC, Yau KW, Dacey DM. Human and macaque pupil responses driven by melanopsin-containing retinal ganglion cells. *Vision Res.* 2007; 47(7):946–54. Epub 2007/02/27. <https://doi.org/10.1016/j.visres.2006.12.015> PMID: 17320141.
11. Zaidi FH, Hull JT, Peirson SN, Wulff K, Aeschbach D, Gooley JJ, et al. Short-wavelength light sensitivity of circadian, pupillary, and visual awareness in humans lacking an outer retina. *Curr Biol.* 2007; 17(24):2122–8. Epub 2007/12/18. <https://doi.org/10.1016/j.cub.2007.11.034> PMID: 18082405.
12. Mure LS, Cornut PL, Rieux C, Drouyer E, Denis P, Gronfier C, et al. Melanopsin bistability: a fly's eye technology in the human retina. *PLoS One.* 2009; 4(6):e5991. Epub 2009/06/25. <https://doi.org/10.1371/journal.pone.0005991> PMID: 19551136.
13. McDougal DH, Gamlin PD. The influence of intrinsically-photosensitive retinal ganglion cells on the spectral sensitivity and response dynamics of the human pupillary light reflex. *Vision Res.* 2010; 50(1):72–87. Epub 2009/10/24. <https://doi.org/10.1016/j.visres.2009.10.012> PMID: 19850061.
14. Gooley JJ, Ho Mien I, St Hilaire MA, Yeo SC, Chua EC, van Reen E, et al. Melanopsin and rod-cone photoreceptors play different roles in mediating pupillary light responses during exposure to continuous light in humans. *J Neurosci.* 2012; 32(41):14242–53. Epub 2012/10/12. <https://doi.org/10.1523/JNEUROSCI.1321-12.2012> PMID: 23055493.
15. Spitschan M. Photoreceptor inputs to pupil control. *J Vis.* 2019; 19(9):5. Epub 2019/08/16. <https://doi.org/10.1167/19.9.5> PMID: 31415056.
16. Zandi B, Klages J, Khanh TQ. Prediction accuracy of L- and M-cone based human pupil light models. *Sci Rep.* 2020; 10(1):10988. Epub 2020/07/06. <https://doi.org/10.1038/s41598-020-67593-3> PMID: 32620793.
17. Rao F, Chan AHS, Zhu XF. Effects of photopic and circopic illumination on steady state pupil sizes. *Vision Res.* 2017; 137:24–8. Epub 2017/07/10. <https://doi.org/10.1016/j.visres.2017.02.010> PMID: 28688906.

18. Lucas RJ, Peirson SN, Berson DM, Brown TM, Cooper HM, Czeisler CA, et al. Measuring and using light in the melanopsin age. *Trends Neurosci.* 2014; 37(1):1–9. Epub 2013/11/30. <https://doi.org/10.1016/j.tins.2013.10.004> PMID: 24287308.
19. Kelbsch C, Strasser T, Chen Y, Feigl B, Gamlin PD, Kardon R, et al. Standards in Pupillography. *Front Neurol.* 2019; 10:129. Epub 2019/03/12. <https://doi.org/10.3389/fneur.2019.00129> PMID: 30853933.
20. Rukmini AV, Milea D, Gooley JJ. Chromatic pupillometry methods for assessing photoreceptor health in retinal and optic nerve diseases. *Front Neurol.* 2019; 10:76. Epub 2019/02/28. <https://doi.org/10.3389/fneur.2019.00076> PMID: 30809186.
21. Park JC, Moura AL, Raza AS, Rhee DW, Kardon RH, Hood DC. Toward a clinical protocol for assessing rod, cone, and melanopsin contributions to the human pupil response. *Invest Ophthalmol Vis Sci.* 2011; 52(9):6624–35. Epub 2011/07/12. <https://doi.org/10.1167/iovs.11-7586> PMID: 21743008.
22. Bonmati-Carrion MA, Hild K, Isherwood C, Sweeney SJ, Revell VL, Skene DJ, et al. Relationship between human pupillary light reflex and circadian system status. *PLoS One.* 2016; 11(9):e0162476. Epub 2016/09/17. <https://doi.org/10.1371/journal.pone.0162476> PMID: 27636197.
23. Adhikari P, Zele AJ, Thomas R, Feigl B. Quadrant field pupillometry detects melanopsin dysfunction in glaucoma suspects and early glaucoma. *Sci Rep.* 2016; 6:33373. Epub 2016/09/14. <https://doi.org/10.1038/srep33373> PMID: 27622679.
24. Feigl B, Zele AJ, Fader SM, Howes AN, Hughes CE, Jones KA, et al. The post-illumination pupil response of melanopsin-expressing intrinsically photosensitive retinal ganglion cells in diabetes. *Acta Ophthalmol.* 2012; 90(3):e230–4. Epub 2011/09/03. <https://doi.org/10.1111/j.1755-3768.2011.02226.x> PMID: 21883986.
25. Chellappa SL, Steiner R, Oelhafen P, Cajochen C. Sex differences in light sensitivity impact on brightness perception, vigilant attention and sleep in humans. *Sci Rep.* 2017; 7(1):14215. Epub 2017/10/29. <https://doi.org/10.1038/s41598-017-13973-1> PMID: 29079823.
26. Munch M, Leon L, Crippa SV, Kawasaki A. Circadian and wake-dependent effects on the pupil light reflex in response to narrow-bandwidth light pulses. *Invest Ophthalmol Vis Sci.* 2012; 53(8):4546–55. Epub 2012/06/07. <https://doi.org/10.1167/iovs.12-9494> PMID: 22669721.
27. Zele AJ, Feigl B, Smith SS, Markwell EL. The circadian response of intrinsically photosensitive retinal ganglion cells. *PLoS One.* 2011; 6(3):e17860. Epub 2011/03/23. <https://doi.org/10.1371/journal.pone.0017860> PMID: 21423755.
28. Griefahn B, Künemund C, Bröde P, Mehnert P. The validity of a German version of the Morningness-Eveningness-Questionnaire developed by Horne and Ostberg. *Somnologie (Berl).* 2001; 5(2):71–80. <https://doi.org/10.1046/j.1439-054X.2001.01149.x>
29. Lanthony P. The desaturated panel D-15. *Doc Ophthalmol.* 1978; 46(1):185–9. <https://doi.org/10.1007/BF00174107> PMID: 310382
30. Ishihara S. Tests for colour-blindness. Tokyo: Kanehara Shuppan Co., LTD.; 1972.
31. Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci.* 2011; 22(11):1359–66. Epub 2011/10/19. <https://doi.org/10.1177/0956797611417632> PMID: 22006061.
32. CIE. CIE S 026/E:2018. CIE system for metrology of optical radiation for ipRGC-influenced responses of light. In: CIE, editor. CIE S 026/E:2018. Vienna: CIE; 2018.
33. Dacey DM, Liao H-W, Peterson BB, Robinson FR, Smith VC, Pokorny J, et al. Melanopsin-expressing ganglion cells in primate retina signal colour and irradiance and project to the LGN. *Nature.* 2005; 433(7027):749–54. <https://doi.org/10.1038/nature03387> PMID: 15716953
34. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
35. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.* 2012; 9(7):671–5. <https://doi.org/10.1038/nmeth.2089> PMID: 22930834
36. Daneault V, Vandewalle G, Hebert M, Teikari P, Mure LS, Doyon J, et al. Does pupil constriction under blue and green monochromatic light exposure change with age? *J Biol Rhythms.* 2012; 27(3):257–64. Epub 2012/06/02. <https://doi.org/10.1177/0748730412441172> PMID: 22653894.
37. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem.* 1964; 36(8):1627–39. <https://doi.org/10.1021/ac60214a047>
38. Iadanza E, Goretti F, Sorelli M, Melillo P, Pecchia L, Simonelli F, et al. Automatic detection of genetic diseases in pediatric age using pupillometry. *IEEE Access.* 2020; 8:34949–61. <https://doi.org/10.1109/access.2020.2973747>
39. Supporting Information to this publication in the Open Science Framework Database [Internet]. 2021.

40. CIE. ISO/CIE 11664–1:2019(E). Colorimetry—Part 1: CIE standard colorimetric observers. ISO/CIE 11664–1:2019(E). Vienna: CIE; 2019.
41. CIE. CIE 165:2005. CIE 10 Degree photopic photometric observer. CIE 165:2005. Vienna: CIE; 2005.
42. Lang D. Human Centric Lighting. Daylight-Related Metrics as prerequisite for assessment of light quality and for lighting design 2016. <http://lightingforpeople.eu/2016/wp-content/uploads/2016/10/Day-light-related-metrics-for-human-centric-lighting.pdf>.
43. CIE. ISO 23539:2005(E) / CIE S 010/E:2004. Photometry—the CIE system of physical photometry. ISO 23539:2005(E) / CIE S 010/E:2004. Vienna: CIE; 2004.
44. DIN. DIN 5031–2:1982:03. Optical radiation physics and illuminating engineering; evaluation of radiation by different detectors. DIN 5031–2:1982:03. Berlin: Beuth publishing company; 1982.
45. DIN. DIN 5031–3:1982:03. Optical radiation physics and illuminating engineering; quantities, symbols and units of illuminating engineering. DIN 5031–3:1982:03. Berlin: Beuth publishing company; 1982.
46. CIE. CIE TN 004:2016. The use of terms and units in photometry—implementation of the CIE system for mesopic photometry. CIE TN 004:2016. Vienna: CIE; 2016.
47. Lund D, Marshall J, Mellerio J, Okuno T, Schulmeister K, Sliney D, et al., editors. A computerized approach to transmission and absorption characteristics of the human eye. CIE; 2012.
48. Wood S. Generalized Additive Models: An Introduction with R, Second Edition. Boca Raton: CRC Press, Taylor & Francis Group; 2017.
49. Hastie T, Tibshirani R. Generalized additive models. *Statist Sci.* 1986; 1(3):297–310. <https://doi.org/10.1214/ss/1177013604>
50. Sóskuthy M. Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. arXiv: 1703.05339v1 [stat.AP]; 2017.
51. Hastie T, Tibshirani R. Generalized additive models for medical research. *Stat Methods Med Res.* 1995; 4(3):187–96. <https://doi.org/10.1177/096228029500400302> PMID: 8548102.
52. Ringer RVC, Allison M.; Larson, Adam M.; Loschky, Lester C. Crowding in natural scenes. *Iperception.* 2021.
53. Zauner J, Plischke H, Stijnen H, Schwarz UT, Strasburger H. Influence of common lighting conditions and time-of-day on the effort-related cardiac response. *PLoS One.* 2020; 15(10). <https://doi.org/10.1371/journal.pone.0239553> PMID: 33027252
54. Simpson GL. Modelling palaeoecological time series using generalised additive models. *Front Ecol Evol.* 2018; 6. <https://doi.org/10.3389/fevo.2018.00149>
55. Pedersen EJ, Miller DL, Simpson GL, Ross N. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ.* 2019; 7:e6876. Epub 2019/06/11. <https://doi.org/10.7717/peerj.6876> PMID: 31179172.
56. Wood SN, Pya N, Säfken B. Smoothing parameter and model selection for general smooth models. *J Am Stat Assoc.* 2017; 111(516):1548–63. <https://doi.org/10.1080/01621459.2016.1180986>
57. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw.* 2015; 67(1). <https://doi.org/10.18637/jss.v067.i01>
58. Xu R. Measuring explained variation in linear mixed effects models. *Stat Med.* 2003; 22(22):3527–41. <https://doi.org/10.1002/sim.1572> PMID: 14601017
59. Veitch JA, Knoop M. CIE TN 011:2020. What to document and report in studies of ipRGC-influenced responses to light. CIE TN 011:2020. Vienna: CIE; 2020.
60. Joyce DS, Feigl B, Zele AJ. The effects of short-term light adaptation on the human post-illumination pupil response. *Invest Ophthalmol Vis Sci.* 2016; 57(13):5672–80. Epub 2016/10/27. <https://doi.org/10.1167/iovs.16-19934> PMID: 27784072.
61. Zele AJ, Cao D. Vision under mesopic and scotopic illumination. *Front Psychol.* 2014; 5:1594. Epub 2015/02/07. <https://doi.org/10.3389/fpsyg.2014.01594> PMID: 25657632.
62. Spitschan M, Jain S, Brainard DH, Aguirre GK. Opponent melanopsin and S-cone signals in the human pupillary light response. *Proc Natl Acad Sci U S A.* 2014; 111(43):15568–72. Epub 2014/10/15. <https://doi.org/10.1073/pnas.1400942111> PMID: 25313040.
63. Adhikari P, Zele AJ, Feigl B. The post-illumination pupil response (PIPR). *Invest Ophthalmol Vis Sci.* 2015; 56(6):3838–49. Epub 2015/06/13. <https://doi.org/10.1167/iovs.14-16233> PMID: 26066752.
64. Lee SK, Sonoda T, Schmidt TM. M1 intrinsically photosensitive retinal ganglion cells integrate rod and melanopsin inputs to signal in low light. *Cell Rep.* 2019; 29(11):3349–55 e2. Epub 2019/12/12. <https://doi.org/10.1016/j.celrep.2019.11.024> PMID: 31825819.
65. Patterson SS, Kuchenbecker JA, Anderson JR, Neitz M, Neitz J. A color vision circuit for non-image-forming vision in the primate retina. *Curr Biol.* 2020. Epub 2020/02/23. <https://doi.org/10.1016/j.cub.2020.01.040> PMID: 32084404.

66. Mouland JW, Martial F, Watson A, Lucas RJ, Brown TM. Cones support alignment to an inconsistent world by suppressing mouse circadian responses to the blue colors associated with twilight. *Curr Biol*. 2019; 29(24):4260–7 e4. Epub 2019/12/18. <https://doi.org/10.1016/j.cub.2019.10.028> PMID: 31846668.
67. Brown TM, Thapan K, Arendt J, Revell VL, Skene DJ. S-cone contribution to the acute melatonin suppression response in humans. *J Pineal Res*. 2021: e12719. Epub 2021/01/30. <https://doi.org/10.1111/jpi.12719> PMID: 33512714.
68. Spitschan M, Lazar R, Yetik E, Cajochen C. No evidence for an S cone contribution to acute neuroendocrine and alerting responses to light. *Curr Biol*. 2019; 29(24):R1297–R8. Epub 2019/12/18. <https://doi.org/10.1016/j.cub.2019.11.031> PMID: 31846672.
69. Guido ME, Marchese NA, Rios MN, Morera LP, Diaz NM, Garbarino-Pico E, et al. Non-visual opsins and novel photo-detectors in the vertebrate inner retina mediate light responses within the blue spectrum region. *Cell Mol Neurobiol*. 2020. Epub 2020/11/25. <https://doi.org/10.1007/s10571-020-00997-x> PMID: 33231827.
70. Morad Y, Lemberg H, Yofe N, Dagan Y. Pupillography as an objective indicator of fatigue. *Curr Eye Res*. 2000; 21(1):535–42. Epub 2000/10/18. PMID: 11035533.
71. McHill AW, Sano A, Hilditch CJ, Barger LK, Czeisler CA, Picard R, et al. Robust stability of melatonin circadian phase, sleep metrics, and chronotype across months in young adults living in real-world settings. *J Pineal Res*. 2021:e12720. Epub 2021/02/02. <https://doi.org/10.1111/jpi.12720> PMID: 33523499.
72. Okimura K, Nakane Y, Nishiwaki-Ohkawa T, Yoshimura T. Photoperiodic regulation of dopamine signaling regulates seasonal changes in retinal photosensitivity in mice. *Sci Rep*. 2021; 11(1):1843. Epub 2021/01/21. <https://doi.org/10.1038/s41598-021-81540-w> PMID: 33469071.
73. DIJK D-J, DUFFY JF, CZEISLER CA. Circadian and sleep/wake dependent aspects of subjective alertness and cognitive performance. *Journal of Sleep Research*. 1992; 1(2):112–7. <https://doi.org/10.1111/j.1365-2869.1992.tb00021.x> PMID: 10607036
74. Webler FS, Spitschan M, Foster RG, Andersen M, Peirson SN. What is the 'spectral diet' of humans? *Curr Opin Behav Sci*. 2019; 30:80–6. Epub 2019/08/23. <https://doi.org/10.1016/j.cobeha.2019.06.006> PMID: 31431907.
75. Rufiange M, Beaulieu C, Lachapelle P, Dumont M. Circadian light sensitivity and rate of retinal dark adaptation in indoor and outdoor workers. *J Biol Rhythms*. 2007; 22(5):454–7. Epub 2007/09/19. <https://doi.org/10.1177/0748730407305375> PMID: 17876066.
76. Roenneberg T. Having Trouble Typing? What on Earth Is Chronotype? *J Biol Rhythms*. 2015; 30(6):487–91. Epub 2015/10/09. <https://doi.org/10.1177/0748730415603835> PMID: 26446872.
77. Roenneberg T. What is chronotype? *Sleep Biol Rhythms*. 2012; 10(2):75–6.
78. Martinez-Lozano N, Barraco GM, Rios R, Ruiz MJ, Tvarijonaviciute A, Fardy P, et al. Evening types have social jet lag and metabolic alterations in school-age children. *Sci Rep*. 2020; 10(1):16747. Epub 2020/10/09. <https://doi.org/10.1038/s41598-020-73297-5> PMID: 33028896.
79. Goldstein D, Hahn CS, Hasher L, Wiprzycka UJ, Zelazo PD. Time of day, intellectual performance, and behavioral problems in morning versus evening type adolescents: is there a synchrony effect? *Pers Individ Dif*. 2007; 42(3):431–40. <https://doi.org/10.1016/j.paid.2006.07.008> PMID: 17268574.
80. Facer-Childs E, Brandstaetter R. The impact of circadian phenotype and time since awakening on diurnal performance in athletes. *Curr Biol*. 2015; 25(4):518–22. Epub 2015/02/03. <https://doi.org/10.1016/j.cub.2014.12.036> PMID: 25639241.
81. Khalsa SBS, Jewett ME, Cajochen C, Czeisler CA. A phase response curve to single bright light pulses in human subjects. *J Physiol*. 2003; 549(3):945–52. <https://doi.org/10.1113/jphysiol.2003.040477> PMID: 12717008
82. Prayag AS, Munch M, Aeschbach D, Chellappa SL, Gronfier C. Light modulation of human clocks, wake, and sleep. *Clocks Sleep*. 2019; 1(1):193–208. Epub 2020/04/29. <https://doi.org/10.3390/clockssleep1010017> PMID: 32342043.
83. Roenneberg T, Mellow M. Entrainment of the human circadian clock. *Cold Spring Harb Symp Quant Biol*. 2007; 72:293–9. Epub 2008/04/19. <https://doi.org/10.1101/sqb.2007.72.043> PMID: 18419286.
84. Danilenko KV, Plisov IL, Cooper HM, Wirz-Justice A, Hebert M. Human cone light sensitivity and melatonin rhythms following 24-hour continuous illumination. *Chronobiol Int*. 2011; 28(5):407–14. Epub 2011/07/05. <https://doi.org/10.3109/07420528.2011.567425> PMID: 21721856.
85. Besharse J, Dunis D. Methoxyindoles and photoreceptor metabolism: activation of rod shedding. *Sci*. 1983; 219(4590):1341–3. <https://doi.org/10.1126/science.6828862> PMID: 6828862
86. Stone RA, Pardue MT, Iuvone PM, Khurana TS. Pharmacology of myopia and potential role for intrinsic retinal circadian rhythms. *Exp Eye Res*. 2013; 114:35–47. Epub 2013/01/15. <https://doi.org/10.1016/j.exer.2013.01.001> PMID: 23313151.



87. Prigge CL, Yeh PT, Liou NF, Lee CC, You SF, Liu LL, et al. M1 ipRGCs influence visual function through retrograde signaling in the retina. *J Neurosci*. 2016; 36(27):7184–97. Epub 2016/07/08. <https://doi.org/10.1523/JNEUROSCI.3500-15.2016> PMID: 27383593.
88. Welbourne LE, Morland AB, Wade AR. Human colour perception changes between seasons. *Curr Biol*. 2015; 25(15):R646–7. Epub 2015/08/05. <https://doi.org/10.1016/j.cub.2015.06.030> PMID: 26241135.
89. Fischer D, Lombardi DA, Marucci-Wellman H, Roenneberg T. Chronotypes in the US—Influence of age and sex. *PLoS One*. 2017; 12(6):e0178782. Epub 2017/06/22. <https://doi.org/10.1371/journal.pone.0178782> PMID: 28636610.
90. Rukmini AV, Milea D, Aung T, Gooley JJ. Pupillary responses to short-wavelength light are preserved in aging. *Sci Rep*. 2017; 7:43832. Epub 2017/03/08. <https://doi.org/10.1038/srep43832> PMID: 28266650.
91. Winn B, Whitaker D, Elliott DB, Phillips NJ. Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science*. 1994; 35(3):1132–7. PMID: 8125724
92. Adhikari P, Pearson CA, Anderson AM, Zele AJ, Feigl B. Effect of age and refractive error on the melatonin mediated post-illumination pupil response (PIPR). *Sci Rep*. 2015; 5:17610. Epub 2015/12/02. <https://doi.org/10.1038/srep17610> PMID: 26620343.