Visual Neuroscience

Pulses of Melanopsin-Directed Contrast Produce Highly Reproducible Pupil Responses That Are Insensitive to a Change in Background Radiance

Harrison McAdams,¹ Aleksandra Igdalova,¹ Manuel Spitschan,² David H. Brainard,³ and Geoffrey K. Aguirre¹

¹Department of Neurology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, United States
²Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom
³Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania, United States

PURPOSE. To measure the pupil response to pulses of melanopsin-directed contrast, and compare this response to those evoked by cone-directed contrast and spectrally narrowband stimuli.

METHODS. Three-second unipolar pulses were used to elicit pupil responses in human subjects across three sessions. Thirty subjects were studied in session 1, and most returned for sessions 2 and 3. The stimuli of primary interest were “silent substitution” cone- and melanopsin-directed modulations. Red and blue narrowband pulses delivered using the post-illumination pupil response (PIPR) paradigm were also studied. Sessions 1 and 2 were identical, whereas session 3 involved modulations around higher radiance backgrounds. The pupil responses were fit by a model whose parameters described response amplitude and temporal shape.

RESULTS. Group average pupil responses for all stimuli overlapped extensively across sessions 1 and 2, indicating high reproducibility. Model fits indicate that the response to melanopsin-directed contrast is prolonged relative to that elicited by cone-directed contrast. The group average cone- and melanopsin-directed pupil responses from session 3 were highly similar to those from sessions 1 and 2, suggesting that these responses are insensitive to background radiance over the range studied. The increase in radiance enhanced persistent pupil constriction to blue light.

CONCLUSIONS. The group average pupil response to stimuli designed through silent substitution provides a reliable probe of the function of a melanopsin-mediated system in humans. As disruption of the melanopsin system may relate to clinical pathology, the reproducibility of response suggests that silent substitution pupillometry can test if melanopsin signals differ between clinical groups.

Keywords: melanopsin, ipRGCs, PIPR, pupillometry

Melanopsin is a photopigment found within the intrinsically photosensitive retinal ganglion cells (ipRGCs; Fig. 1a). Although they represent a small fraction (~1%–3%) of the total retinal ganglion cell population,¹⁻³ ipRGCs are critical for entrainment of circadian rhythm,³⁻⁶ aversive responses to light,⁷ light-induced lacrimation,⁸ and control of pupil diameter.⁹⁻¹¹ Disruption of these reflexive visual functions is seen in many clinical conditions, leading to the speculation that dysfunction in the melanopsin system is responsible.⁷,¹²⁻¹⁷ Consequently, there is interest in measuring, in humans, a signal that reflects melanopsin function and testing if this signal varies between groups.

The post-illumination pupil response (PIPR) paradigm is one method to assess melanopsin function in humans.¹¹⁻¹⁸,¹⁹ The PIPR paradigm exploits the differing spectral sensitivities of the melanopsin photopigment and the cone-based luminance mechanism: the medium- and long-wavelength cones (M and L), which are the primary input to the luminance mechanism, are more sensitive to light of longer wavelengths (“red”), whereas melanopsin sensitivity is greatest in the short-wavelength (“blue”) range. The PIPR paradigm measures the response of the pupil to pulses of narrowband blue and red light presented against steady dark backgrounds. Particular attention is paid to the behavior of the pupil at relatively delayed time periods, including after stimulus offset (i.e., “post-illumination”), when melanopsin is found to exert greater and more sustained influence over pupil size relative to the cones.¹¹⁻²⁰ In fact, this relative persistence of the melanopsin response as compared with the more rapidly adapting cone response is a key property of signaling within ipRGCs. PIPR measurements have been made in numerous clinical conditions, including multiple sclerosis, Parkinson’s disease, idiopathic intracranial hypertension, traumatic brain injury, glaucoma, diabetes, retinitis pigmentosa, Leber’s hereditary optic neuropathy, Smith-Magenis syndrome, and depression.²¹⁻³²
Although relatively simple to deploy and measure, interpretation of the PIPR as a melanopsin-specific signal is less straightforward. Because the blue stimulus is presented against a dark background, the pupil response will include a rod contribution.\textsuperscript{33,34} Blue light also drives S cones, which, like melanopsin, produce delayed and sustained pupil responses.\textsuperscript{35} Although there is convincing evidence that sustained pupil constriction can be produced by melanopsin alone,\textsuperscript{11} cones may also contribute (perhaps via the ipRGCs) to a sustained response.\textsuperscript{36,37} Therefore, although the PIPR response reflects (perhaps overwhelmingly) the contribution of melanopsin signals, it cannot be concluded that differences between clinical populations in PIPR responses are attributable solely to the melanopsin system.

Silent substitution spectral modulations\textsuperscript{38} provide an alternate approach to the study of the melanopsin contribution to the human pupil response.\textsuperscript{35,39–44} Light spectra are tailored to modulate the response of one or more targeted photoreceptor mechanisms (e.g., melanopsin, while holding the response of the remaining photoreceptor mechanisms (e.g., L, M, and S cones) constant. Subjects first adapt to a background light spectrum. When the silent substitution modulation is presented around that background, the subsequent response is attributable to the targeted photoreceptor(s). Here, we measured the temporal properties and reliability of the across-subject average pupil response to pulses of melanopsin contrast delivered via silent substitution. We compared the response to melanopsin stimulation to that evoked by cone-directed contrast that was silent for melanopsin, and by narrowband PIPR stimuli. To anticipate, we find that the silent substitution approach produces a highly reproducible measure of melanopsin-driven pupil response that is insensitive to a change in background radiance.

![Figure 1](image-url)
These experiments were the subject of preregistration documents. The preregistered protocol was followed (with small exceptions, see Methods) in subject recruitment, screening, exclusion, stimulus validation, and pupil data preprocessing. The analyses described in the preregistration examined the reliability of between-subject differences in response. We found relatively low reliability and present those results in the Supplementary Materials (Supplementary Figs S5, S6). We focus here on population level analyses that were not preregistered.

METHODS

Subjects

Subjects were recruited from the community of students and staff at the University of Pennsylvania. Exclusion criteria for enrollment included a prior history of glaucoma or a negative reaction to pupil-dilating eye drops. During an initial screening session, subjects were also excluded for abnormal color vision as determined by the Ishihara plates or visual acuity below 20/40 in each eye as determined using a distance Snellen eye chart. Subjects completed a brief, screening pupilometry session. We excluded at this preliminary stage subjects who were unable to provide high-quality pupil tracking data (details below). Poor data quality was found to result from difficulty suppressing blinks or from poor infrared contrast between the pupil and iris.

A total of 32 subjects were recruited and completed initial screening. Two of these subjects were excluded after screening due to poor data quality (e.g., excessive loss of data points from blinking) as determined by preregistered criteria. Thirty subjects thus successfully completed session 1 and provided data for analysis. These subjects were between 19 and 53 years of age (mean 25.93 ± 4.24 SD). Fourteen subjects identified as male, 15 female, and 1 declined to provide a gender identification. Of this group of 30 subjects, 24 completed an identical second session of testing and 21 completed a third session at higher light levels. The time between participation in session 1 and session 2 was on average 110 days, and between session 1 and session 3 on average 296 days. The study was approved by the Institutional Review Board of the University of Pennsylvania, with all subjects providing informed written consent, and all experiments adhered to the tenets of the Declaration of Helsinki.

When a subject arrived for a session of primary data collection, the right eye was first anesthetized with 0.5% proparacaine and dilated with 1% tropicamide ophthalmic solution. Subjects then had their right eye dark adapted by wearing swimming goggles with the right eye obscured while sitting in a dark room for 20 minutes. In an attempt to minimize variation in circular cycle across sessions, testing for sessions 2 and 3 started within 3 hours of the time of day when the same subject started session 1.

Stimuli

The experiments used two classes of stimuli: (1) silent substitution spectral modulations that targeted either the melanopsin photopigment or the cone-mediated luminance postreceptor mechanism; (2) narrowband blue and red stimuli designed to elicit the PIPR.

The silent substitution stimuli were a subset of those used in a prior report, and full details of their generation may be found there. Briefly, we used the method of silent substitution together with a digital light synthesis engine (OneLight Spectra, Vancouver, BC, Canada) to stimulate targeted photoreceptors. The device produces stimulus spectra as mixtures of 56 independent primaries (~16 nm full width at half maximum [FWHM]) under digital control, and can modulate between these spectra at 256 Hz. Details regarding the device, stimulus generation, and estimates of precision have been previously reported. Our estimates of photoreceptor spectral sensitivities were as previously described with those for the cones based on the field size and age-dependent International Commission on Illumination (CIE) physiological cone fundamentals. The estimates account for subject age. Pupil size (which was fixed at 6-mm diameter through the use of an artificial pupil), and our field size of 27.5 degrees. Although the standard specifies fundamentals only for field sizes up to 10 degrees, we obtained the 27.5-degree estimates by extrapolating the formula from the standard routines in the open-source Psychophysics Toolbox. Separate background and modulation spectra were identified to provide nominal 400% Weber contrast on melanopsin while silencing the cones for the melanopsin-directed background/modulation pair (Mel), and 400% contrast on each of the L-, M-, and S-cone classes while silencing melanopsin for the luminance-directed modulation/background pair (LMS) (Fig. 1c). The xy chromaticities of the background spectra for the Mel and LMS stimuli were similar (Mel: ~0.56, ~0.40; LMS: ~0.58, ~0.38). The background for Mel and LMS pulses were nominally rod-saturating (~110 photopic cd/m² or 3.10 log scotopic trolands for Mel and ~40 photopic cd/m² or 2.99 log scotopic trolands for LMS for sessions 1 and 2; ~270 photopic cd/m² or 3.59 log scotopic trolands and ~90 photopic cd/m² or 3.46 log scotopic trolands for session 3). The xy chromaticities and photopic luminances reported above were calculated using the proposed XYZ functions associated with the CIE 2006 10-degree cone fundamentals (http://www.cvrl.org). The modulations did not explicitly silence rods or penumbral cones.

The PIPR stimuli consisted of narrowband pulses of blue (475 ± 25-nm peak ± Gaussian FWHM) and red (625 ± 25-nm) light (Fig. 1d). These stimuli were each designed to produce 12.30 log quanta-cm⁻² sec⁻¹ retinal irradiance for sessions 1 and 2, and 12.85 log quanta-cm⁻² sec⁻¹ for session 3, in a manner that accounted for differences in lens density due to subject age. These stimuli were presented against a dim background (~0.5 cd/m² for the first two sessions, ~1 cd/m² for the third session). The irradiance of the PIPR stimuli was limited by the gamut of the device at short wavelengths, and the requirement to match the retinal irradiance of the red and blue stimuli. Background light levels were the minimum possible with our apparatus, as some light is emitted by the light engine even when all primaries are set to their minimum level.

Due to imperfections in device control, the actual stimuli presented differed in photoreceptor contrast and irradiance from their nominal designed values. Before and after each subject’s measurement session, spectroradiometric validation measurements of the background and modulation spectra were obtained. From these, we calculated the actual contrast on targeted and nominally silenced photoreceptors for that subject (using age-based photoreceptor sensitivities) for the silent substitution stimuli, as well as the retinal irradiance of the PIPR stimuli. Following our preregistered protocol, we excluded data for a given session if the postexperiment validation measurements showed that the silent substitution stimuli were of insufficient quality. Specifically, if the contrast on the targeted postreceptor mechanism (Mel or LMS) was less than 350% (as compared with the nominal 400%), or if contrast on an ostensibly silenced postreceptor mechanism (Mel, LMS, L–M, or S) was greater than 20%. Data from five sessions were discarded (and subsequently recollected) as a consequence of this procedure. We did not evaluate the PIPR
stimuli for the purposes of data exclusion. Supplementary Table S1 provides the results of the stimulus validations for all subjects, sessions, and stimuli. These calculations do not account for the biological variability in individual photoreceptor spectral sensitivity that can produce further departures from nominal stimulus contrasts. 8

Three-second pulses of spectral change were presented during individual trials of 17-seconds duration (Fig. 1e). During each trial, a transition from the background to the stimulation spectrum (Mel, LMS, blue, or red) would occur starting at either 0, 1, or 2 seconds after trial onset (randomized uniformly across trials); this jitter was designed to reduce the ability of the subject to anticipate the moment of stimulus onset. The transition from the background to the stimulation spectrum, and the return to background, was smoothed by a 500-ms half-cosine window. The half-cosine windowing of the stimulus was designed to minimize perception of a Purkinje tree percept in the melanopsin-directed stimulus. 8

Each session consisted of three blocks of stimuli: PIPR (consisting of both red and blue stimuli counterbalanced in order within subject), LMS, and Mel, in this fixed order. At the start of each block, the subject adapted to the background spectrum for 4.5 minutes. The block consisted of 24, 17-second trials. Within each block, after every six trials, participants were invited to take a break before resuming the experiment. During the break they could lift their head from the chin rest. The duration of each break was determined by the subject, and was less than a few minutes. Light adaptation was not maintained during the break. Before continuing with the experiment, subjects re-adapted to the background spectrum for 30 seconds, whether or not they took a break.

Stimuli were presented through a custom-made eyepiece with a circular, uniform field of 27.5° diameter and the central 5° diameter obscured (Fig. 1f). The central area of the stimulus was obscured to minimize stimulation within the macula, where macular pigment alters the spectral properties of the stimulus arriving at the photoreceptors. Subjects viewed the field through a 6-mm-diameter artificial pupil and were asked to maintain fixation on the center of the obscured central region.

**Pupillometry**

Pupil diameter was measured using an infrared video pupillometry system (Video Eye Tracker; Cambridge Research Systems Ltd., Rochester, UK), sampled at 50 Hz. Following acquisition, the raw measured pupil response was adjusted in time to account for the stimulus-onset time within each trial and normalized by the baseline pupil size for that trial (with baseline size taken as the mean pupil diameter for 1 second before stimulus onset). Data points in the resulting response for which the velocity of constriction or dilation exceeded 2500% change/s were rejected and replaced via linear interpolation. The responses across trials were averaged.

Pupillometry data were excluded from analysis on the basis of the number of rejected data points. Trials containing 10% or more rejected data points were deemed incomplete and excluded from the average; if more than 75% of the trials for a given stimulus type were excluded, then the entire session was judged to be incomplete and the subject was either restudied or excluded, following our preregistered procedure. Additionally, if more than 50% of trials across all stimulus types were excluded, then the subject was either restudied or excluded. Data from four sessions were discarded for this reason; two of these subjects were restudied.

As noted briefly under the “Subjects” section above, screening pupillometry was also performed before primary data collection to exclude subjects for whom good-quality pupil tracking data could not be obtained. In a screening session, subjects were presented two sets of six trials of the PIPR stimuli. Subjects with four or more incomplete trials assessed by the same criterion above were excused from the experiment. Two subjects were excused from the study in this manner.

**Analysis**

We fit the pupil response for each stimulus and subject using a three-component temporal model (Fig. 2a). 41 The stimulus profile passes through the model and, under the control of six parameters, is transformed into a predicted pupil response. The six parameters include two time constants that influence the shape of each component, three gain parameters that adjust the scaling of each component, and one onset delay parameter that shifts the entire modeled response in time. The transient component captures the initial peak of pupil constriction, the sustained component tracks the shape of the stimulus profile, and the persistent component describes the slow dilation of the pupil back to baseline. Each component has an amplitude parameter. The shape of the components are under the control of two temporal parameters. The $\tau_\gamma$ parameter controls the rate of onset and width of all components. The $\tau_x$ parameter controls the rate of exponential decay of the persistent component. The three components are summed to create the model response, which is then temporally shifted in time by the overall delay parameter. We fit this model to the average response for each subject for each stimulus condition. In analyzing group differences of model parameters, the median value was used, as parameters were not normally distributed across subjects.

Model fits were performed using MATLAB’s (MathWorks, Natick, MA, USA) fmincon function. Fits were initialized from six different starting positions and the fit with the highest proportion variance explained ($R^2$) was retained. Additionally, bounds were placed on each parameter, as informed by an initial inspection of the data. The bounds of the $\tau_\gamma$ were different for responses elicited through silent substitution and PIPR stimuli. Specifically, the upper boundary of $\tau_\gamma$ for fits to responses elicited by PIPR stimuli was greater than that for fits to responses elicited through silent substitution to reflect the generally wider shape of these responses. This choice improved the quality of fits to each stimulus type. As we were interested exclusively in comparisons within a stimulus type (LMS versus Mel, red versus blue), the differing parameter boundaries would not influence any subsequent conclusions. We also performed additional analyses in which we locked and freed different sets of parameters as part of control tests. These procedures are described in the Supplementary Materials (Supplementary Fig. S3).

To test for significance of observed group differences of metrics derived from our model, we used label permutation. For a given group comparison, we took the observed metric aggregated across all trials for a given stimulus type for each subject and randomly assigned each metric to the correct stimulus label or the opposite stimulus label. After performing this for all subjects, we computed the median difference. We performed this simulation 1,000,000 times, and asked the percentage of simulations in which the simulated median difference is more extreme than the observed median difference.

**Preregistration of Studies**

Our studies (composed of three sessions of data collection) were the subject of preregistration documents (https://osf.io/...
The preregistered protocol dictated subject recruitment, screening, data exclusion, and stimulus validation. Session 1 was designed to test if we could measure the melanopsin-mediated pupil response to silent substitution and PIPR stimuli in individuals. Data collection for session 1 commenced in September 2016. An addendum (https://osf.io/hyj89/) detailed an improvement in our approach to generating stimuli that accurately described stimulus production for both the initial and subsequent subjects; this document is dated September 2016 but was not uploaded until October 2016. A January of 2017 addendum clarified an ambiguity in our original description of the stimulus validation procedure (https://osf.io/btr3q/).

Session 2 was designed to determine if the magnitude of pupil response to melanopsin stimulation was a reliable individual subject difference (https://osf.io/z2vj7/). Session 3 repeated the measurements at a higher light level in an attempt to evoke a larger response to the PIPR stimuli and to further test the reliability of any individual differences in pupil response (https://osf.io/anguy/).

Our original motivation for these studies was to measure individual differences in pupil response. We ultimately determined that this test was limited by within-session measurement noise. Therefore, this paper focuses on comparisons at the group level. In keeping with our preregistered protocols, however, we provide in the supplementary material the results of individual subject analyses (Supplementary Figs. S5, S6).

There was ambiguity in our initial protocol regarding the interpretation of postexperimental stimulus validations. Five validation measurements were made after each experimental session. Our preregistration initially failed to delineate how to interpret all five validation values; our procedure was later updated to include an additional addendum (https://osf.io/bg76w/) and an annotation (https://osf.io/bg76w/).

Reproducible Pupil Response to Melanopsin Contrast

FIGURE 2. A three-component model used to fit the group average pupil responses. (a) Within-subject average evoked responses to each stimulus type was subjected to nonlinear fitting with a six-parameter, three-component model. The model was designed to capture the visually apparent and temporally separated components of the evoked pupil response. The elements of the model are not intended to directly correspond to any particular biological mechanism. The input to the model was the stimulus profile (black). An additional input vector, representing the rate of stimulus change at onset, was created by differentiating the stimulus profile and retaining the positive elements. These three vectors were then subjected to convolution operations composed of a gamma and exponential decay function (blue), each under the control of a single time-constant parameter (t\text{\_gamma} and t\text{\_exponential}). The resulting three components (red) were normalized to have unit area, and then subjected to multiplicative scaling by a gain parameter applied to each component (g_{\text{transient}}, g_{\text{sustained}}, and g_{\text{persistent}}). The scaled components were summed to produce the modeled response (gray), which was temporally shifted (t\text{\_delay}). This caption and the corresponding panel are adapted from Figure S9 of Spitschan et al.41 (b) The model fit, computed from the median response parameter across all 30 subjects from session 1, is plotted in dotted lines on top of the group average response from session 1. The gray inset shows each model component of the fit (transient, sustained, and persistent in most to least saturated color).
clarified to specify that data from a session would be excluded if the median value across all five postexperiment validation values was larger than the cutoff criterion. Data from one session were discarded and recollected based on an initial interpretation of the validation procedure in which a single validation measurement that exceeded criterion led to data rejection. Following the clarification of our procedure to use the median validation measurement, data from four subsequent sessions were discarded and recollected because of stimulus quality.

Availability of Data and Analysis Code
Data will be available via figshare on publication. Analysis code that operates on the raw data and produces the results and figures may be found at: https://github.com/gkaguirrelab/pupilPIPRAnalysis.

RESULTS
In each of 30 subjects we measured consensual pupil responses in the left eye to spectral modulations presented to the pharmacologically dilated right eye (Fig. 1b). Two of the modulations targeted the postreceptoral luminance or melanopsin pathway using a silent substitution spectral exchange (Fig. 1c, left), and two of the modulations were narrowband red or blue increments typical of PIPR studies (Fig. 1c, right). We presented these spectral modulations as half-cosine windowed, 3-second pulses (Fig. 1d) on a spatially uniform field, except for masking of the central 5° of visual angle to minimize stimulation of the macula (Fig. 1e). We recorded the ensuing pupil response for each of many trials in 30 subjects (session 1) and a subset of these subjects in sessions 2 (24 subjects) and 3 (21 subjects). We derived the average pupil response for each subject across the 24 trials presented in a

![Figure 3](image-url)
session. In supplementary analyses (Supplementary Figs. S1, S2) we examined the dependence of pupil size and pupil response on trial order.

Silent Substitution and PIPR Stimuli Elicit Highly Reproducible Pupil Responses at the Group Level

We first examined the form of group (averaged over subjects) pupil responses to pulsed spectral modulations designed to selectively target the cones or melanopsin (Fig. 3a, top row). We measured pupil responses during the 13 seconds that followed the onset of a 3-second stimulus pulse, and expressed pupil size as the percentage change in diameter relative to the prestimulus period. For our silent substitution stimuli, which were equated in contrast, the LMS-mediated pupil response was of overall larger amplitude than that evoked by the melanopsin-directed stimulus. The responses also differed in their shape, with the offset of the stimulus producing a more rapid dilation for LMS stimulation as compared with melanopsin stimulation.

The red and blue PIPR stimuli also produced pupil constriction. These stimuli were equated in retinal irradiance but the amplitude of pupil constriction was smaller in response to the red stimulus as compared with the blue stimulus. The shape of these responses also differed subtly, as the pupil began to dilate during the red stimulus, while the constriction in response to the blue stimulus continued to increase during stimulus presentation.

The SEM across subjects was quite small relative to the amplitude of response. Although this might suggest that the measurements would be reproducible in this group, it is possible that variation in subject state (e.g., due to seasonal or circadian changes) or drift in our apparatus would reduce reproducibility across sessions. We tested for reproducibility by repeating the measurements during session 2 in 24 of the 30 subjects between 54 and 175 days later (Fig. 3a, bottom row). The amplitude, shape, and within-session SEM measured in session 2 was quite similar to that measured in session 1. Figure 3b presents the group average from session 2 plotted directly on top of that from session 1 for each stimulus condition. The reproducibility of the pupil response to all stimuli is evident, both in amplitude (max absolute difference in amplitude of group-averaged responses: 1.47% for LMS, 1.64% for Mel, 2.24% for red, 1.74%, for blue), and in shape (Pearson correlation coefficient of the session 1 group average with the session 2 group average: r = 0.999 for LMS, r = 0.995 for Mel, r = 0.998 for red, r = 0.999 for blue).

The Melanopsin Response Is More Persistent Than the Cone Response

Melanopsin-driven activation of ipRGCs results in notably prolonged responses. Here we asked if a difference in the temporal profile of the pupil response to cone and melanopsin stimulation is apparent at the group level. We fit the data from each subject with a three-component model of the pupil response (Fig. 2a). The model has amplitude parameters for transient, sustained, and persistent components, as well as three temporal parameters that specify the overall timing and influence the shape of the components. Figure 2b illustrates the model fits for the data from session 1. The fit line is given by the median of the model parameters across subjects. The amplitude and shape of the three model components are shown inset in each panel. After combining the data from sessions 1 and 2 for those subjects studied twice, we tested for differences in the amplitude and temporal parameters evoked by the different stimuli.

The transient, sustained, and persistent components of the model reflect different temporal domains. The persistent component captures the slow return to baseline of the pupil response following the offset of the stimulus. We considered that stimulation of the ipRGCs might produce pupil responses with a relatively enhanced persistent component. For each subject for each stimulus type, we computed the proportion of the total pupil response area made up of the persistent component (Fig. 4a). Across subjects, the median pupil response to LMS stimulation had 50% of its total response area...
fit by the persistent component. In contrast, the response to melanopsin stimulation was 76% persistent ($P = 0.0015$ established by permutation of stimulus labels). This difference reflects primarily a larger sustained component in the pupil response to luminance; the absolute response area of the persistent component was similar for the cone- and melanopsin-driven responses (Supplementary Table S2). Unexpectedly, for the PIPR stimuli, the persistent component was larger in response to the red as compared with the blue stimulus (median "percent persistent" for red: 65%; for blue: 58%; $P = 0.00047$ by label permutation).

We considered that the temporal profile of the persistent response, as opposed to its magnitude alone, would reflect the influence of melanopsin. The model parameter $\tau_{\text{exponential}}$ influences the rate at which the persistent component of pupil diameter dilates back to baseline following stimulus offset. We tested if this time constant differed in the responses to the stimulus types. Consistent with the expected properties of the ipRGCs, the melanopsin-driven response had a slower return to baseline as compared with the LMS-driven response (Fig. 4b, median $\tau_{\text{exponential}}$ for LMS: 3.46 seconds; for Mel: 6.90 seconds; $P = 0.00047$ by label permutation). This slower return to baseline was also observed for the response to the blue stimulus as compared with the red stimulus (red: 4.89 seconds; blue: 7.50 seconds; $P = 0.0039$ by label permutation). Supplementary Table 2 contains the amplitude and temporal parameters for all conditions and stimuli.

A property of our analysis is that the temporal parameters are allowed to vary between the compared stimulus conditions to best fit the data. It is therefore possible that observed differences in the $\tau_{\text{exponential}}$ or "percent persistent" measurements arise as a consequence of differences in other model parameters. To evaluate this possibility, we re-ran the analyses holding the other temporal parameters fixed between the two compared stimulus conditions. This analysis revealed very similar results (Supplementary Fig. S3).

Silent Substitution and PIPR Methods Are Differently Sensitive to Stimulus Radiance

We considered the possibility that the pupil response evoked by the silent substitution stimuli would be relatively insensitive to the overall spectral power of the stimuli, as long as the contrast was held constant. For session 3, we modified our apparatus to increase the radiance of all stimuli. Although the background luminance of the silent substitution stimuli more than doubled (mean background luminance for LMS increased from $\sim 40 \text{ cd/m}^2$ to $\sim 90 \text{ cd/m}^2$ or $2.99 \log \text{ scotopic trolands}$ to $3.46 \log \text{ scotopic trolands}$; for Mel increased from $\sim 110 \text{ cd/m}^2$ to $\sim 270 \text{ cd/m}^2$ or $3.10 \log \text{ scotopic trolands}$ to $3.59 \log \text{ scotopic trolands}$), the calculated LMS and melanopsin contrast remained the same. For the PIPR stimuli, the nominal intensity of the spectral pulse increased from 12.30 to 12.85 log quanta cm$^{-2}$ sec$^{-1}$, and the background luminance increased from 0.5 cd/m$^2$ to 0.9 cd/m$^2$.

We then repeated the pupil measurements in 21 of the 30 subjects between 238 and 352 days after their initial enrollment (session 3). Figure 5 presents the group average response collapsed across the first two sessions, compared with the pupil response measured in session 3. For the LMS
and Mel stimuli, the average group response was essentially unchanged (max absolute difference in amplitude of group-averaged responses: 1.10% for LMS, 1.27% for Mel; Pearson correlation of the evoked response between session 1/2 and session 3: LMS, \( r = 0.999 \); Mel, \( r = 0.994 \)). This high degree of reproducibility suggests that the pupil response to the silent substitution stimuli is insensitive to this change in absolute light intensity and instead reflects stimulus contrast.

In distinction, the increase in the radiance of the PIPR stimuli produced a larger amplitude of pupil response (max absolute difference in amplitude of group-averaged responses: 3.60% for red, 4.89% for blue). Many studies that use the PIPR stimuli attempt to isolate the melanopsin-specific component by taking the difference of the blue and red responses. Figure 6 presents the difference in pupil response evoked by the red and blue stimuli at the two radiance levels. This PIPR effect, especially at the later time points, grows in magnitude from sessions 1 and 2 to session 3 (Fig. 6). We quantified the PIPR effect as the difference in the total response area of the model fits to blue and red stimuli for each subject, for each session. The median PIPR was larger in session 3 as stimulus radiance was increased (sessions 1/2 median PIPR: 35% change * s; session 3 median PIPR: 74% change * s; \( P = 0.0002 \) by label permutation).

**DISCUSSION**

We find that pulsed spectral modulations that target the cones and melanopsin evoke distinctive pupil responses. At a group level, the average responses to these silent substitution stimuli are highly reliable. Consistent with the known temporal properties of the ipRGCs, the response to melanopsin-directed as compared with cone-directed stimulation features a relatively larger persistent response that returns to baseline more slowly.

Our findings indicate the feasibility of using pupilometry with silent substitution stimuli to test for group differences in cone and melanopsin physiology. As compared with the PIPR stimuli, the silent substitution approach more directly targets and isolates the melanopsin and cone systems. Further, the highly reproducible responses seen at the group level indicate that differences between groups should be detected with good statistical power. Indeed, the extent to which this group average signal is reliable can be seen in the highly similar responses elicited from a different cohort of subjects using the same stimuli as part of a previous study.41 (Supplementary Fig. S4).

We applied a model to the temporal profile of pupil responses to derive amplitude and timing parameters. This model accounts well for the form of response to both silent substitution and PIPR stimuli. Although we use “percent persistent” to describe differences between the cone- and melanopsin-driven pupil responses, we find that all stimuli evoke some degree of persistent response. This observation is consistent with prior work, both in previous PIPR studies that show that the red stimulus evokes persistent pupil constriction, as well as neurophysiologic studies that show ipRGCs generate persistent firing from nonmelanopsin inputs.37 We examined as well the \( t_{\text{exponential}} \) timing parameter of our model fits. The melanopsin-directed and PIPR blue stimuli produced responses with greater \( t_{\text{exponential}} \) values as compared with their cone-directed and PIPR red counterparts. Therefore, although all stimulus types evoked some amount of persistent pupil response, slower resolution of this response was seen for the stimuli thought to drive melanopsin. We anticipate that the temporal model may be used to test for differences in the amplitude and temporal properties of melanopsin-driven responses in clinical populations.

Although these results suggest that a key feature of the melanopsin response is its persistence, the delay to response onset was also found to differ between melanopsin- and cone-driven pupil responses. The relatively longer delay to response onset for the melanopsin-driven response is broadly consistent with prior work, although the absolute magnitude is less than might be expected.5,20,52,53 Comparing the latency of neural
responses within the ipRGCs to that within pupil responses is not straightforward. Neural response latency is dependent on stimulus intensity, and can be as short as several hundred milliseconds for melanopsin-driven ipRGC activation. Further, nonlinearities in the conversion of retinal ganglion cell signals to pupil responses would complicate the interpretation of absolute latency differences between cone- and melanopsin-driven signals.

An original motivation for our study was to examine individual differences in the pupil response. Although average responses at the group level were highly reliable, we found that there was relatively poor reproducibility for individual subjects (Supplementary Fig. S5). We examined the reproducibility of total pupil response amplitude across subjects. Although there was a reasonable correlation of this measure between sessions 1 and 2, these responses did not correlate with the measurements from session 3. Our results do not reject the possibility that there are in fact reliable individual differences in the pupil response. Simulations suggest that within-session measurement noise could have obscured a true individual difference effect. Analysis of individual subject data also failed to show a relationship between individual differences in melanopsin function as elicited through the silent substitution and PIPR approaches (Supplementary Fig. S6). In future studies, increasing the number of trials and improving pupillometry quality could reduce within-session measurement error and perhaps reveal reproducible individual differences in response.

In session 3, we examined the effect of a multiplicative increase in stimulus intensity. This manipulation increased the radiance of both the stimulus and the background. For the silent substitution stimuli that targeted either the cones or melanopsin, this change in stimulus intensity did not alter the pupil response. Although retinal irradiance was increased in session 3, the contrast of the silent substitution modulations remained constant at 40%. Therefore, within this stimulus regime, the pupil response to silent substitution stimuli appears to be best characterized in terms of the photoreceptor contrast of the modulation.

These results also allow us to discount the possibility of inadvertent rod stimulation by the melanopsin-directed stimulus. The spectral sensitivity functions of melanopsin and rhodopsin overlap. Consequently, the melanopsin-directed silent substitution stimulus has substantial calculated contrast (~320%) on the rod photoreceptors. Because the stimulus background is in the photopic range, we generally assume that the rods are saturated, and thus this rod contrast does not contribute to the observed pupil response. This assumption, however, may be challenged by recent work that finds rods can signal above their nominal saturation threshold. It is therefore reassuring to observe in the current study that the pupil response is unchanged with the increased stimulus intensity used in session 3. If there were a substantial rod contribution to the pupil response measured to the melanopsin-directed stimulus in sessions 1 and 2, we would expect that this contribution would become smaller at the higher background level. The equivalence of the pupil response suggests that any rod signals are minimal under these conditions. There remain other mechanisms through which rods could impact our measured pupil responses, including through the possibility of light scatter onto rods in the obscured parfoveal region or in the periphery beyond our 27.5-degree field. Prior work suggests, however, that any rod signaling produced through scattered light would have a small and transient effect on the measured response.

Conversely, the PIPR evoked by the chromatic stimuli was enhanced by the increase in stimulus intensity in session 3. Similar to the silent substitution stimuli, the photoreceptor contrast produced by the red and blue stimuli is in principle unchanged in session 3. However, small imperfections in the control of the dim background light levels used for the PIPR stimuli could produce substantial changes in stimulus contrast. Although our stimulus measurements indicate fairly consistent calculated contrast between the experimental sessions (Supplementary Table S1), actual variation in the contrasts produced by the PIPR stimuli remains a possible explanation for the enhanced responses to the PIPR stimuli seen in session 3. It is also possible, however, that the increased pupil response to the PIPR stimuli is a real effect of the change in stimulus intensity. When stimuli are presented against dark backgrounds, changes in intensity can lead to substantial changes in rod activation, which could then alter the response.

We note that our implementation of the PIPR paradigm differs from that used in many other studies, due to the particular nature of our apparatus. For example, the change in stimulus intensity examined in session 3 increased both the stimulus and background light levels. This is unlike previous studies of the dependence of the PIPR on intensity, in which the background presumably was held fixed across changes in the intensity of the chromatic pulses. Our apparatus also imposes gamut limits that restrict how dark we can make the background and how intense we can make the chromatic pulses. In addition, many PIPR studies make use of a Ganzfeld dome and thus provide a greater spatial extent of stimulation than used here. These difference likely account for the smaller magnitude of PIPR that we obtain in comparison with other studies.

Acknowledgments

Supported by National Institutes of Health Grant R01EY024681: melanopsin and cone signals in human visual processing.

Disclosure: H. McAdams, None; A. Igdalova, None; M. Spitschan, P; D.H. Brainard, P; G.K. Aguirre, P

References