Standardized and calibrated light stimuli via head-mounted displays for investigating the non-visual effects of light

Maydel Fernandez-Alonso[1] & Manuel Spitschan[1,2,3]

1 Translational Sensory & Circadian Neuroscience, Max Planck Institute for Biological Cybernetics, Tübingen, Germany
2 Chronobiology & Health, TUM School of Medicine and Health, Technical University of Munich, Munich, Germany
3 TUM Institute for Advanced Study (TUM-IAS), Technical University of Munich, Germany

Abstract

Light influences human physiology profoundly, affecting the circadian clock and suppressing the endogenous hormone melatonin. Experimental studies often employ either homogenous full-field stimulation, or overhead illumination, which are hard to standardize across studies and laboratories. Here, we present a novel technique to examine non-visual responses to light using virtual-reality (VR) head-mounted displays (HMDs) for delivering standardized and calibrated light stimuli to observers in a reproducible and controlled fashion. We find that VR HMDs are well-suited for delivering standardized stimuli defined in luminance and across time, with excellent properties up to 20 Hz. We examine melatonin suppression to continuous luminance-defined light stimuli in a sample of healthy participants (n=31, mean±SD age: 27.4±5.6), and find robust melatonin suppression in 27 out of 31 participants (87% of the sample). Our findings demonstrate that VR HMDs are well-suited for studying the mechanisms underlying human non-visual photoreception in a reproducible and standardized fashion.

Introduction

The non-visual effects of light

Light plays a fundamental role in allowing us to visually perceive the world around us, but it also exerts a profound influence on other physiological and behavioural processes in humans such as circadian entrainment, regulation of hormone secretion, and modulation of alertness and mood (Blume et al., 2019; Brown et al., 2022). These non-visual effects are predominantly mediated by a specialized subset of retinal cells containing the photopigment melanopsin, the intrinsically-photosensitive retinal ganglion cells (ipRGCs) (Berson et al., 2002; Dacey et al., 2005; Do, 2019; Do et al., 2009; Hattar et al., 2002; Lucas et al., 2014; Spitschan, 2019). Over the past two decades, numerous studies have demonstrated that light exposure at specific intensities and wavelengths can acutely suppress the hormone melatonin, affect alertness levels, and shift the timing of the circadian clock (Brown, 2020; Brown et al., 2022). As such, precisely controlled light interventions have become a key tool for researchers investigating a range of outcomes, from basic circadian biology to clinical applications in sleep and mood disorders.

Current methods for light stimulation and limitations

In circadian and sleep research, the most common approaches to deliver light stimuli include overhead room illumination, light boxes or ganzfeld dome stimulators (Spitschan et al., 2023), often equipped with multi-channel LED light sources that allow researchers to vary the intensity levels and spectral compositions of the illumination (Barrionuevo et al., 2022; Brainard et al., 2001; Thapan et al., 2001). More recently, some studies have also incorporated digital displays as a light source in evening studies to deliver a controlled light intervention (Blume et al., 2024; Schöllhorn et al., 2023).

Despite the widespread adoption of these techniques, they carry several limitations. When using overhead room illumination, participants' head movements, eye movements and changes in posture can cause variability in the amount of light actually reaching the retina (Spitschan, 2021; Thibos et al., 2018), leading to uncontrolled differences in exposure that may confound experimental outcomes. Second, in more controlled setups, participants are required to remain in carefully controlled conditions while limiting considerably their posture and head movements, reducing ecological validity and the comfort of the participant. The need for carefully controlled laboratory conditions also discards the possibility to conduct at-home studies. Third, the nature of these setups often imposes limitations in cost, laboratory space

or both, having the consequence that data collection is limited to one participant per evening, significantly slowing the rate at which studies can be completed.

Head-mounted displays for controlled stimulus presentation

The past decade has witnessed a rapid development of virtual and augmented reality technologies. From their early development in the mid-2010s, head-mounted displays (HMDs) have experienced exponential growth, with a very wide variety of affordable and more capable devices available today. While most HMDs have been oriented towards the consumer market, researchers have been quick to take advantage of these new technologies as powerful instruments for conducting behavioural research with humans. Virtual reality (VR) creates an immersive three-dimensional environment that can be fully controlled by the experimenter. Improvements in display hardware, such as better resolution and greater field of view, ensure a more precise stimulus presentation and improve immersion. Furthermore, advancements in rendering technologies and the increased accessibility of software platforms to create realistic virtual environments, has made developing experiments in VR increasingly common (Creem-Regehr et al., 2023).

Beyond recreating real environments, HMDs can also help solve some of the problems with traditional experimental setups, particularly in the field of visual perception. For a typical screen setup, it is crucial to maintain a fixed distance and viewing angle between the observer and the screen, which is often achieved by highly restricting the participant's posture and head movements with the use of chin rests or similar. Furthermore, it is often important to control that no visual stimulation is presented concurrently, which creates the need for completely dark environments where one can ensure that no other stimulus will appear in the observer's field of view. As a result, such setups are often only feasible to implement in a laboratory, making them impractical for other applications outside this setting.

Since HMDs are strapped to the observer's face, the distance between the eyes and the display remains constant, and their immersive nature ensures that researchers have full control of the stimuli that appear within the field of view of the observer. These advantages have led to HMDs being used for psychophysical experiments (Creem-Regehr et al., 2022; Cwierz et al., 2021; De Gelder et al., 2018; Lepori et al., 2023), clinical studies (Mees et al., 2020; Neugebauer et al., 2021; Tatiyosyan et al., 2020; Wiley et al., 2022), and also for applications outside the lab, such as at-home vision testing (Chia et al., 2023; Hu et al., 2023; Tsapakis et al., 2018).

However, one untapped field where the use of HMDs could be advantageous is in light exposure protocols that investigate the non-image-forming effects of light. By delivering a standardized, calibrated light stimulus that encompasses most of the participant's field of view

and remains constant regardless of changes in posture or head movements, HMDs open new possibilities for exploring the mechanisms underlying the non-visual effects of light. One specific application is the study of binocular integration, as HMDs enable precise, independent control of the light delivered to each eye. This capability, combined with the potential to present complex spatially or temporally modulated stimuli, goes beyond what is currently achievable with overhead room illumination or other existing methods.

Integrated eye tracking technology

In more recent times, advancements in HMDs have included integrated eye-tracking capabilities, which make it possible to precisely monitor gaze direction and pupil size in real-time (Adhanom et al., 2023). This feature is particularly relevant to studies on the non-visual light effects of light, as pupil size is directly modulated by the ipRGCs, and at the same time it directly determines retinal irradiance, influencing the magnitude of physiological responses (Gaddy et al., 1993; Higuchi et al., 2008; Schöllhorn et al., 2024). Moreover, tracking gaze direction can help researchers verify compliance (i.e., confirming that participants have their eyes open and are looking at a stimulus for the intended duration). With eye tracking becoming more accessible, the integration of sophisticated calibration and measurement tools within VR HMDs paves the way for a new generation of controlled, individualized light exposure studies that can be performed inside and outside traditional laboratory settings.

Current study

The purpose of the present study is to develop and validate a novel method for delivering standardized, calibrated light stimuli using VR HMDs. We implement temporal and spectral calibration procedures to characterise the displays in use, and implement an evening light exposure intervention to test their feasibility for studying acute melatonin suppression. Ultimately, this approach has the potential to broaden the methodological possibilities within the field of chronobiology, enabling new avenues of research into the non-visual effects of light.

Materials and Methods

Apparatus: Hardware and software setup

Stimuli were presented using a virtual reality (VR) head-mounted display (HMD), the HTC Vive Pro Eye (HTC, Taiwan). This device integrates a high-resolution binocular display with built-in Tobii eye-tracking technology capable of real-time pupillometry and eye movement

measurements. Stimulus control and presentation were managed using custom code developed in Python, using the licensed program Vizard 7 from WorldViz (Santa Barbara, CA, USA). The VR headset allowed for precise control of light stimuli, ensuring calibrated and reproducible light exposure across participants.

The HTC Vive Pro Eye has dual 3.5-inch Organic Light-Emitting Diode (OLED) displays, one for each eye. The technical specifications given by the manufacturer are:

- Resolution: 1440 x 1600 pixels per eye (2880 x 1600 combined)
- Refresh rate: 90 Hz
- Field of view: 110 degrees

Calibration of head-mounted display

Luminance calibration and gamma correction

A luminance calibration of the HMD was performed to ensure a linear light output. Measurements were taken using a Jeti spectraval 1511-HiRes spectroradiometer (Jena Technische Instrument, Jena, Germany) to measure spectral radiance and luminance for each screen primary (RGB) and for white at various input values (from 0 to 1 in steps of 0.1). These measurements were repeated for both displays (i.e. left and right eye). The measurements were processed and fitted using a power function, such that:

$$L = A \cdot I^{\gamma}$$

where L represents the measured luminance, I is the input, A is a scaling factor, and γ is the gamma. The inverse of this exponent was then used to model and correct the non-linear behaviour of the display, such that:

$$I_c = I \cdot ({}^{1}/_{\gamma})$$

Where I is the desired input (e.g., 0.5 for 50% intensity), γ is the estimated gamma value, and I_c is the corrected input to use. We verified the effectiveness of the gamma correction by re-measuring the luminance of the display after correction and confirming a linear luminance output.

Temporal calibration

A temporal calibration of the HTC Vive Pro Eye display was performed to ensure accurate delivery of the intended temporal light stimuli across a range of frequencies. Brightness measurements were recorded over time using a dual channel PicoScope oscilloscope (Pico Technology, St Neots, UK) connected to two mounted DET36A2 photodiodes from Thorlabs (Bergkirchen, Germany). The photodiodes were each placed in front of the respective eye or display of the VR headset, and were used to capture the light output and transform it into a

voltage signal. The photodiodes were securely held in place in front of the display by a custom-made holder to ensure consistent positioning across all measurements.

The measurements were taken at a sampling rate of 250 kHz for a duration of 20 seconds for each stimulus. The conditions tested included a white constant background and sinusoidal flicker at frequencies spaced logarithmically from 0.01 to 20 Hz (0.01, 0.02, 0.05, 0.1, 0.22, 0.5, 1, 2.24, 5, 10, and 20 Hz). A custom-written code was developed to generate and present these sinusoidal flicker frequencies simultaneously on both eyes of the display.

The captured signals were analysed using the Fast Fourier transform implemented in Python to identify the frequency with the peak amplitude and compare it to the input flicker frequency. Additionally, the amplitude at harmonic frequencies was calculated to assess signal fidelity and distortion. The temporal modulation transfer function of the displays was also derived to characterize the system's ability to reproduce different temporal frequencies accurately.

Melatonin suppression study

Data for this study was collected as part of a larger study on binocular retinal signal integration in melatonin suppression (ISRCTN88955347). In addition to the two light conditions in this analysis, participants completed two additional experimental sessions with light conditions that varied temporally or spatially. The protocol followed was identical in all cases, except for five participants that received the light intervention with pharmacologically dilated pupils.

Participants

A total of 31 participants (age range 20-39 and mean ± SD: 27.4 ± 5.6; 11 males, 20 females) were included in this study. Participants were healthy, non-smoking individuals with no reported sleep disorders, substance abuse, medication or hormonal intake. All participants had normal or corrected-to-normal vision, confirmed through screening tests.

Procedure

Participants adhered to a standardized sleep-wake cycle, monitored using actigraphy (Condor ActTrust2, Condor Instruments, São Paolo, Brazil) and daily sleep diaries, for at least 7 days before their first laboratory session, and throughout the 3-weeks study. Alcohol and caffeine intake were restricted throughout the duration of the study. Participants arrived at the laboratory 5 hours before their habitual bedtime, underwent alcohol and THC screening, and completed the calibration procedure of the VR headset eye-tracking system, including adjustment of the headset position on the head, the distances between the two screens, and gaze tracking calibration.

Each experimental session began 4 hours before habitual bedtime. Participants sat in a dimly lit room (<10 lux) during the first 2 hours. Subsequently, they wore the VR headset to view either the dark or bright light stimulus for 20-minute intervals at a time, for four intervals in total. Experimenters were blinded to which light stimulus was being presented each session. While wearing the VR headset, participants were monitored using the integrated eye-tracker of the VR headset. An automated program monitored the openness of both eyes, and issued an audio warning if eyes were detected to be closed for more than 5 seconds. After 2 hours of light exposure through the VR headset, participants returned to the dimly lit room for 1 hour. Saliva samples were collected every 30 minutes throughout the evening, directly after each light exposure block of 20 minutes. Following the saliva sample collection, participants took a break of 5 minutes. Some participants (n=5) viewed the stimuli with pharmacologically dilated pupils. For this, tropicamide eye drops were administered 30 minutes before the light intervention (Mydriaticum Stulln® UD 5 mg/1 ml, Pharma Stulln, Stulln, Germany).

During each evening session, up to 3 participants took part in the study at once. This was made possible by having three separate experimental bays separated by privacy screens, each equipped with an HTC Vive Pro Eye HMD. The use of VR headsets ensured that participants were not exposed to or influenced by the light stimuli presented to others, as each headset delivered individualized and controlled light stimuli directly to the participant's eyes. This setup eliminated any potential cross-exposure to stray light, and ensured a constant light delivery throughout the intervention periods.

Stimuli

The light stimuli consisted of a constant background displayed uniformly across the field of view. In the dark condition, the intensity was set to zero, while in the bright light condition the intensity was set to 50% (after gamma correction and linearization of the display). A fixation cross was positioned centrally to maintain participants' gaze stability during light exposure.

Design

The study employed a within-subjects design, with each participant exposed to both dim and bright light conditions in separate sessions. Sessions were randomized and counterbalanced across participants. Data collection occurred in a controlled laboratory environment, ensuring consistent ambient conditions.

Data processing and analysis

Saliva samples (>1 mL) were collected using Salivettes (Sarstedt, Nürnberg, Germany) and were scanned and centrifuged immediately after collection. Samples were stored at -20°C and

later analysed for melatonin concentration using enzyme-linked immunoabsorbent assay (ELISA) by NovoLytiX GmbH, Switzerland (MLTN-96, with typical specifications: limit of quantification 0.5 –50 pg/mL, detection limit 0.5 pg/mL, mean intra-assay precision: 7.8%, mean inter-assay precision 10.0%).

The melatonin data were processed in several steps using Python. Initially, sample order was verified based on the timestamps at the time of collection, and misordered samples were corrected (n=1 sample). Missing melatonin concentrations were interpolated using the piecewise cubic hermite interpolating polynomial method, supplemented by forward and backward filling for remaining gaps (n=7 samples).

The data was normalized using a weighted mean difference approach. Pre-light exposure data were used to calculate weighted mean melatonin concentrations for each participant across all conditions and separately for each condition. The difference between the participant's overall mean and the condition-specific mean was then computed. This difference was applied to adjust the melatonin concentrations, ensuring normalization across conditions. This normalization was necessary to account for random or unexplained differences observed in the melatonin profiles between sessions, which occurred even before any light intervention was delivered. Since the first two hours of each session were identical across conditions, any differences in melatonin levels during this period could not be attributed to the light intervention. By normalizing the post-light exposure data using pre-light exposure values, we ensured that the observed effects during light exposure were due to the intervention itself and not influenced by inter-session variability or random fluctuations. However, this normalization did not alter the main findings of the paper. In the supplemental materials (**Figure S1**) we provide a comparison of the raw and the normalized data.

The area under the curve (AUC) for melatonin concentrations during light exposure were computed for each condition and participant using Simpson's rule. Finally, percentage melatonin suppression was calculated relative to the dark condition for each participant as:

$$Suppression\ (\%) = \left(1 - AUC_{bright}/AUC_{dark}\right) \cdot 100$$

The statistical analysis was conducted using the lme4 package in R. We fitted a linear mixed-effects model to evaluate the effects of light condition, pupil dilation, and their interaction on normalized AUC values of melatonin concentrations. In addition to the fixed effects of condition and pupil dilation, a random intercept of participant was included to account for inter-individual variability in AUC values. The model incorporating the maximal random effects structure, including both random intercepts and slopes, was tested but failed to converge.

The model was fitted using restricted maximum likelihood estimation (REML). Parameter estimates and interaction effects were assessed, and 95% confidence intervals for the fixed

effects were derived using a bootstrap method to provide robust interval estimates. Model diagnostics were performed to ensure validity, including residual analysis.

Results

Results of HMD calibration

Luminance calibration and gamma correction

Luminance measurements were taken for both eyes of the HMD and for all individual primaries. As expected, we found non-linear relationship between RGB input values and luminance output, which is standard in digital displays. We fitted the gamma curves to each individual RGB primary and to the white background to determine the correction factor. The results of the luminance measurements and the fitted curves are shown in **Figure 1**.

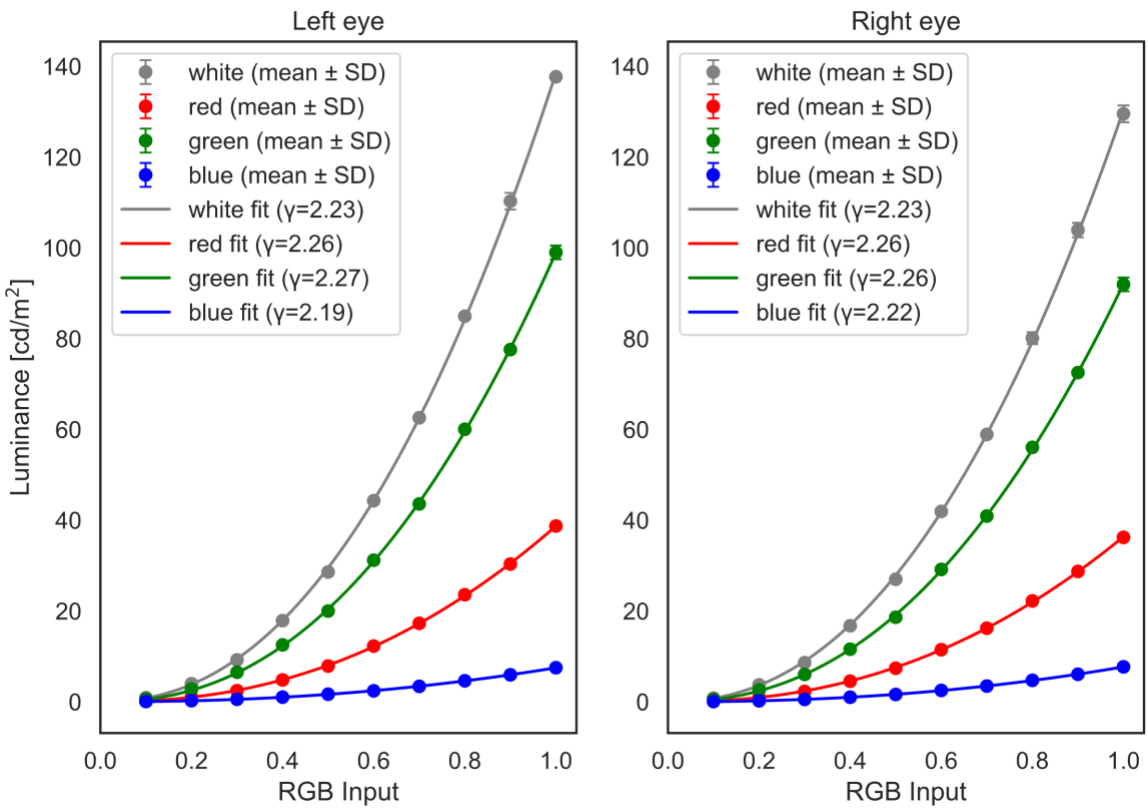

Figure 1. Luminance measurements and fitted gamma functions for each display primary and for white (equal RGB input). The round markers and error bars represent the mean and standard deviation of the measured luminance from three repetitions. The continuous lines represent the fitted power functions, with the estimated gamma values indicated in parenthesis on the legend.

As shown, the resulting gamma values are close to the standard used in most displays, 2.2. However, there are some discrepancies between primaries, particularly the blue colour has a

shallower curve or lower gamma value than the red and green curves. While this difference is small, it could be relevant for studies where precise colour control of the stimulus is important. Since for this study we used a white background as stimuli, we used the corresponding gamma value of 2.225, and calculated the correction factor as the inverse of this. After applying this correction, the luminance output was re-measured and compared against the target linear response. The results are shown in **Figure 2**. These luminance measurements were fitted with a linear regression, with a resulting coefficient of determination (R^2) of 1, indicating that the gamma correction successfully linearized the luminance output of the display.

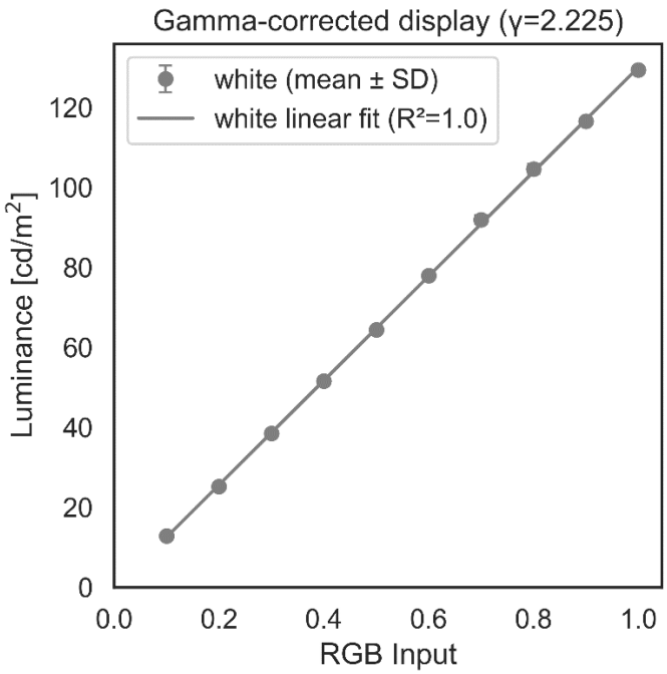

Figure 2. Measured luminance and fitted linear function after gamma correction of the display. The round markers and error bars represent the mean and standard deviation of the measured luminance from three repetitions. The continuous line represents the linear function fitted, with the R2 value of the fit shown in parenthesis on the legend.

Temporal calibration

The luminance of the display was measured for both eyes with high temporal resolution, while displaying sinusoidal flicker at a range of temporal frequencies (from 0.01 to 20 Hz) and for a constant white background. For the latter, the measurements revealed that the display has a constant nearly sinusoidal strobing of the emitted light at 90 Hz, illustrated in **Figure 3**. According to information from the manufacturer, this feature is encoded in the driver of the display and cannot be disabled. Therefore, a peak amplitude at 90 Hz was found for all other measurements taken.

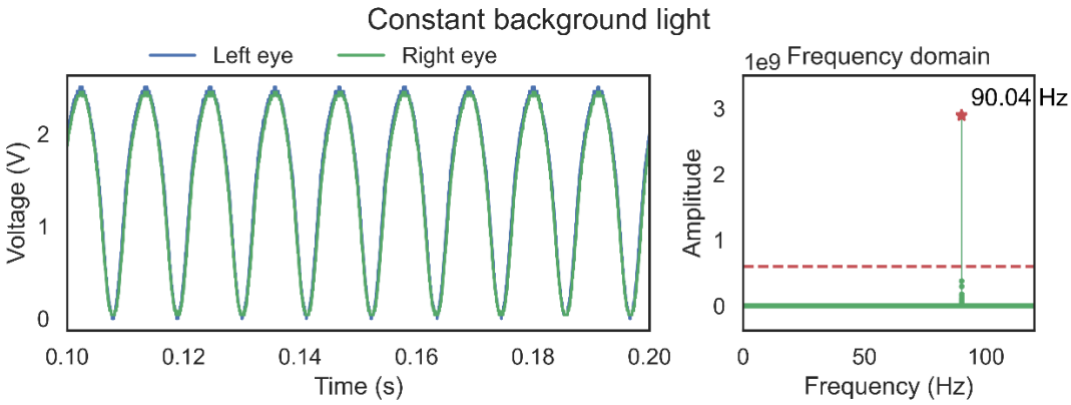

Figure 3. Measured signal in the time domain (left) and in the frequency domain (right) for a constant white background in the left eye and right eye of the display (note that both lines overlap significantly). The dashed red line on the right panel indicates the threshold used to determine the peak amplitude. The frequency with peak amplitude is shown as an annotation (90.04 Hz)

In **Figure 4** we illustrate the results of measurements taken for sinusoidal flicker of 5 Hz. As seen, the strobing of the display at 90 Hz is still present, but the target frequency is displayed with high accuracy and low distortion, as indicated by low power at harmonic frequencies.

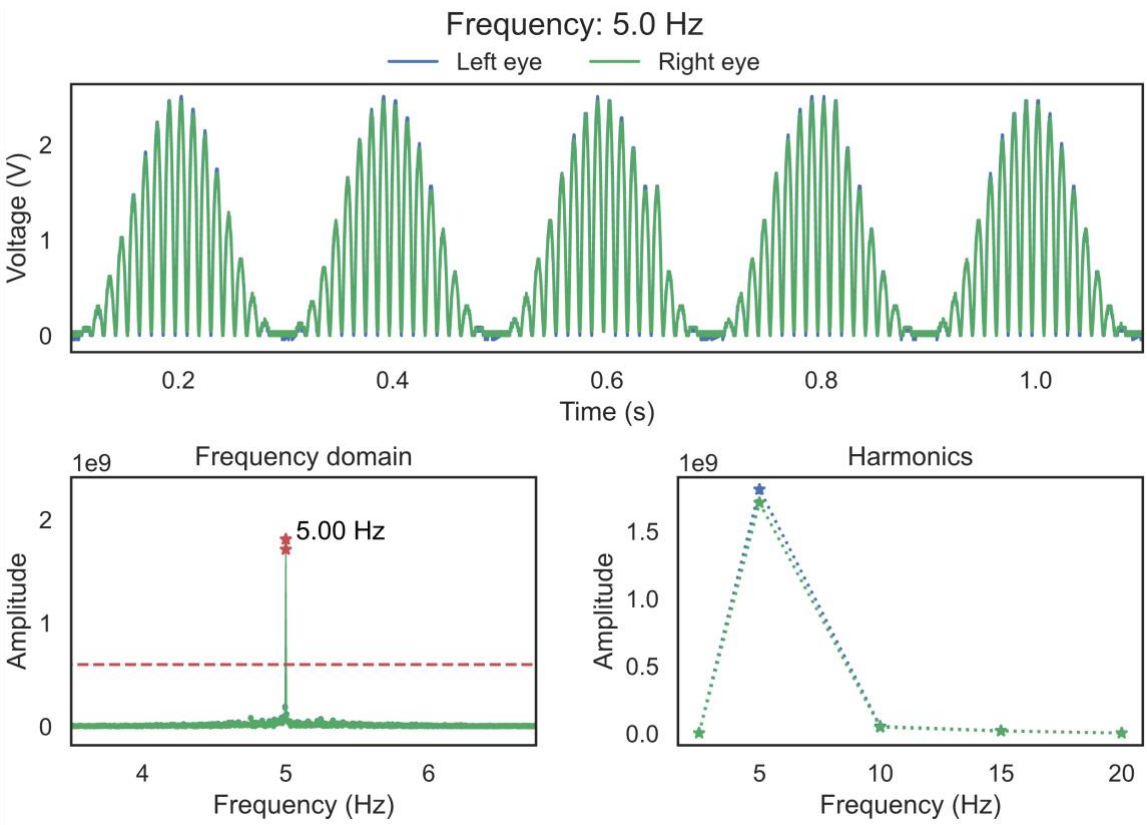

Figure 4. Measured signal for sinusoidal flicker of 5 Hz shown in the left eye and right eye of the display in the time domain (top) and the frequency domain (bottom left). The amplitude at the harmonics of the peak amplitude frequency is shown (bottom right). The dashed red line on the bottom left panel indicates the threshold used to determine the peak amplitude. The frequency with peak amplitude is shown as an annotation on the bottom left figure. Note that the data of both eyes (blue and green) overlap significantly,

Figure *5* summarizes the results of all frequencies tested, displaying temporal transfer function and the amplitude at the harmonics of the target frequencies. Despite the inbuilt strobing, the display was able to reproduce sinusoidal flicker with high accuracy, minimal

distortion and high agreement between both eyes across the measured frequencies, although a significant decrease in amplitude was observed at 20 Hz, reducing the fidelity of the displayed stimuli.

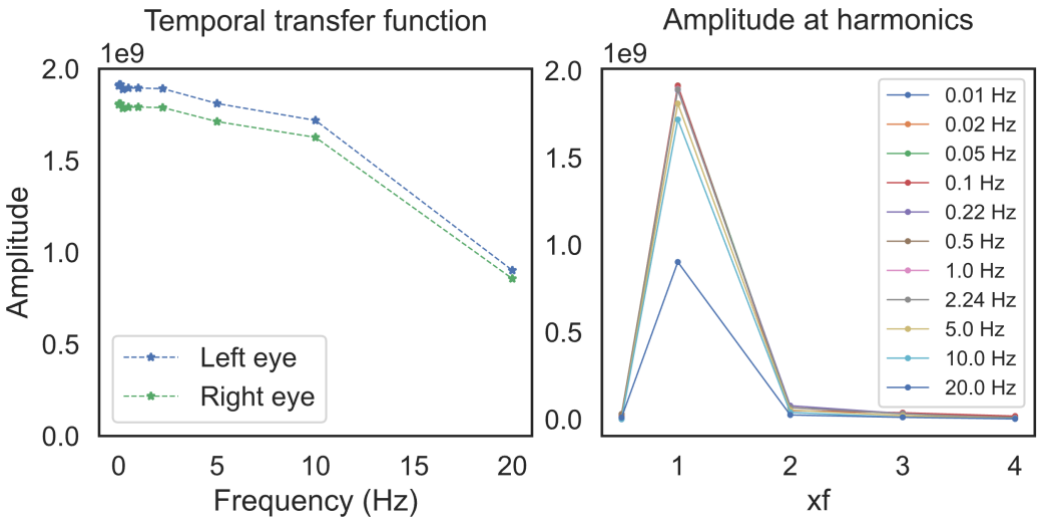

Figure 5. Temporal modulation transfer function of the display (left) and amplitude at the harmonics of each measured frequency.

We also tested higher frequencies such as 25 Hz, but there the interaction between the sinusoidal flicker and the inbuilt strobing severely distorted the waveform, causing high amplitude at higher harmonics. It is possible that this issue could be ameliorated by using a different waveform such as square-wave flicker; although our findings in general indicate that the display is best suited for delivering temporally-modulated stimuli at low to mid-range temporal frequencies.

Results of melatonin suppression study

We investigated the acute melatonin suppression effects of light stimuli delivered with head-mounted virtual-reality displays in 31 healthy adults. We tested two conditions: a dark background, and a constant bright light background of 50% intensity. An example of the results for one participant are shown in **Figure 6**. We applied a normalization procedure to correct for baseline differences in melatonin levels across sessions in the period before light exposure. These differences cannot be explained by the light intervention, since the first 2 hours of every session are spent under the same dim light conditions. Thus, this analysis ensures that post-light exposure comparisons accurately reflect the influence of the light stimuli delivered.

As seen here, the effects this procedure had on the melatonin profiles and the calculated areas under the curve (AUCs) were minimal for most participants and did not have a significant effect on the results presented here. A full comparison of raw and normalized AUCs across all participants is provided in the Supplementary Materials (**Figure S1**).

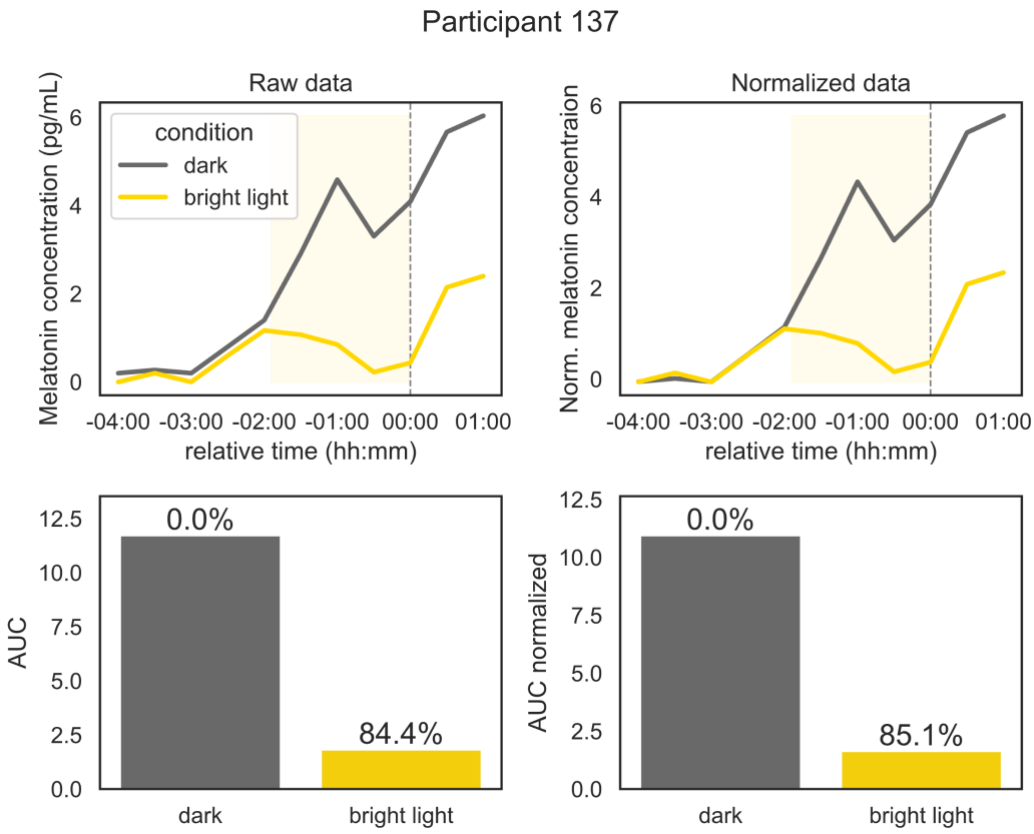

Figure 6. Salivary melatonin concentration as a function of time relative to habitual bedtime (top) and the calculated AUCs during light exposure for each condition (bottom) for participant 137. The two figures on the left show the raw data, while the figures on the right illustrate the effects of the normalization procedure described in the methods. The labels above the bar plots represent the percent of melatonin suppression calculated as a function of the dark condition.

Figure 7 shows the results for all participants in the sample, separated by whether participant's pupils were pharmacologically dilated or not. As shown on the left panel of **Figure 7,** the cumulative melatonin concentration during light exposure is lower in the bright light condition for most participants compared to the dark condition, indicating an acute suppression of melatonin.

To examine this effect statistically, we fitted a linear mixed model on the AUCs, with a fixed effect of condition, pupil dilation, and their interaction, and a random intercept of participant. This means that the baseline AUC in the dark condition was allowed to vary randomly for each participant. A model with the maximal random effects structure (i.e., with a random slope of condition) was also tested, but failed to converge. The results of the model fit are shown in **Table 1**. As seen, there is a significant effect of experimental condition, with the estimated AUC under the bright light condition being significantly lower than under the dark condition by 14.40 (CI 95% -19.19 to -8.95, $p < 0.001$). Regarding pupil dilation, no significant effects were found, neither in the dark condition nor in the bright light condition. This means that participants that took part in the experiment with pupil dilation did not differ significantly from the

participants with natural pupils, and that being exposed to the bright light condition did not have any significant further effects in the AUC in the bright light condition.

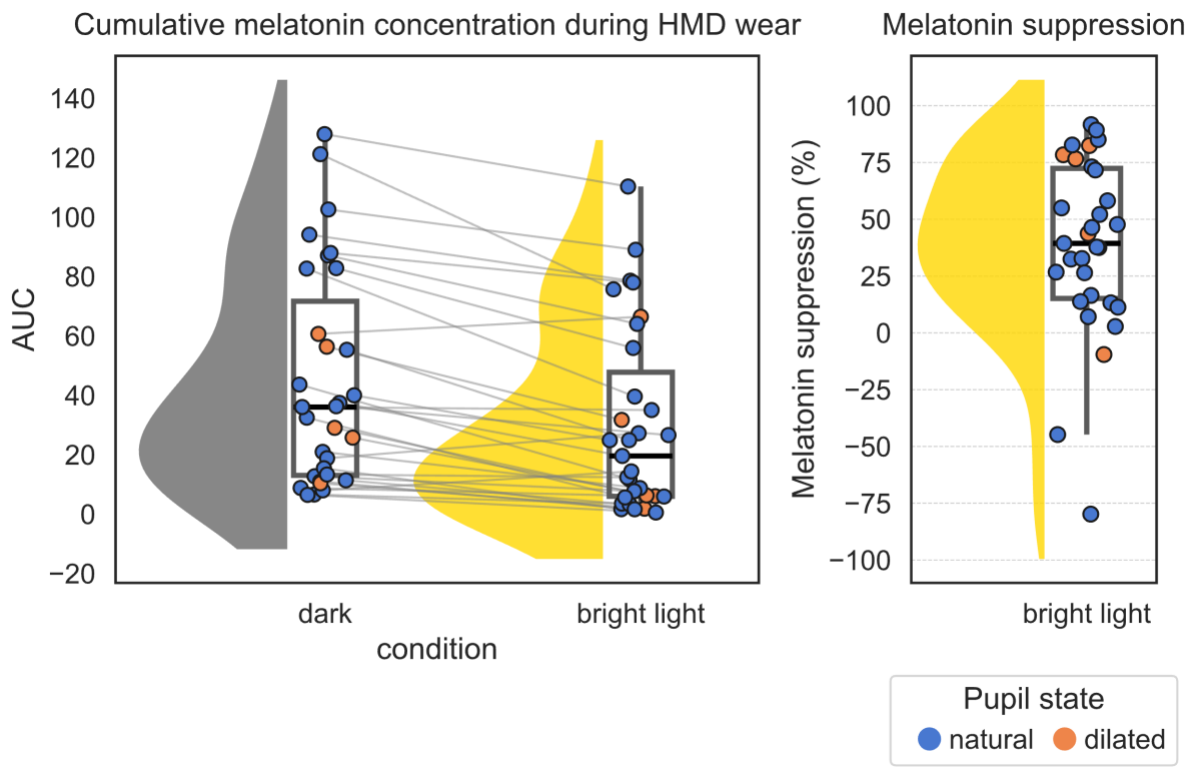

Figure 7. Cumulative melatonin concentration as a function of condition (left) and calculated melatonin suppression for the bright light condition (right). Each marker represents one individual participant. The colour of the markers indicates if the light stimulation was received with a natural or pharmacologically dilated pupil. The continuous grey lines indicate repeated measures, connecting the data points that belong to the same participant.

Table 1. Results of the linear mixed model fit of cumulative melatonin concentration (AUC) during light exposure with fixed effects of condition, pupil state and their interaction, and random intercept of participant. The results shown are the parameter estimates (Est.), the 95% confidence intervals (CI 95%), the standard error (SE), degrees of freedom (df), t-test and p-values, as well as the standard deviation of the random effects (RE SD)

	Est.	**CI 95%**		**SE**	**df**	**t-test**	**p-value**	**RE SD**
Intercept (dark condition, natural pupil)	46.19	32.33	58.86	6.71	31.33	6.89	<0.001	32.86
Bright light	-14.40	-19.19	-8.95	2.64	29.00	-5.46	<0.001	-
Pupil dilation (dark condition)	-9.75	-42.20	23.79	16.71	31.33	-0.58	0.564	-
Pupil dilation * Bright light	0.42	-12.05	14.28	6.57	29.00	0.06	0.949	-
Residual								9.52

To better illustrate what this represents in terms of melatonin suppression, we show in the right panel of **Figure 7** the distribution of the suppression percent in the sample, and we provide descriptive statistics in **Table 2**. On average, melatonin was suppressed by 38.6% in

the bright light condition for all participants. When comparing the subgroups by pupil state, we see that those with dilated pupils had a higher mean melatonin suppression at 54.3%, compared to 35.6 % in the natural pupil group. However, in both groups we see considerable variability in the melatonin suppression effect with some extreme outlier values, which could explain the lack of statistically significant differences between both groups in the previous analysis.

Table 2. Descriptive statistics of melatonin suppression percent for participants that received the light stimuli with dilated or natural pupils, and for all participants.

	Mean	**SD**	**Min**	**25%**	**Median**	**75%**	**Max**
Dilated pupil	54.32 %	38.91 %	-9.53 %	43.73 %	76.58 %	78.38 %	82.46 %
Natural pupil	35.62 %	39.30 %	-79.78 %	14.46 %	37.66 %	57.36 %	91.67 %
All participants	38.64 %	39.21 %	-79.78 %	15.13 %	39.39 %	72.38 %	91.67 %

Discussion

In the present study, we demonstrated that head-mounted virtual-reality displays can be effectively calibrated to deliver controlled light stimuli for chronobiology research. First, we established luminance calibration and gamma correction procedures, verifying that the non-linear relationship between digital input values and luminance output could be successfully linearized. Despite minor discrepancies between primary colours—particularly for blue—our final gamma correction achieved a near-perfect linear response (R^2 = 1). Second, temporal calibration revealed the presence of a built-in 90 Hz strobing in the display driver, which was not user-accessible for deactivation. Nonetheless, the system reproduced sinusoidal flicker at low-to-mid temporal frequencies (up to 20 Hz) with minimal distortion.

We also completed a melatonin suppression study involving 31 healthy adults, and we observed a significant reduction in melatonin levels under a bright light condition (white background at 50% intensity) compared to a dark condition (black background). This suppression reached an average of 38.6% across participants, demonstrating that VR displays can elicit robust acute non-visual light responses. Increasing the brightness of the screen to 100% would be expected to produce an even greater suppression effect (Prayag et al., 2019). Contrary to previous findings (Gaddy et al., 1993), the use of pharmacological pupil dilation did not confer a statistically significant change in melatonin suppression, likely due to substantial inter-individual variability and the small number of subjects that underwent pupil dilation.

Advantages of head mounted displays

Enhanced precision and reduced variability

The calibration procedures demonstrated that VR HMDs can provide high-precision control over both the spectral and temporal properties of the light stimulus. This level of consistency surpasses what is usually achievable with traditional setups, where factors such as viewing angle and room reflections can introduce unwanted variability in retinal illumination (Spitschan, 2021). By delivering light directly through an HMD, each participant receives the same stimulus, regardless of posture or external conditions.

Automation and eye-tracking integration

An additional benefit of this system lies in the integration of real-time eye tracking. By monitoring participants with an automated program, the experimenters were able to confirm that participants were indeed viewing the stimulus, thereby reducing the need for constant supervision. In general, the programmability of HMDs supported a high degree of automation in this study. We employed custom software to deliver audio instructions throughout the experiment and precisely schedule steps like light exposure and saliva sampling. The use of audio instructions and the integration of an external response box, allowed participants to advance through the experimental protocol on their own by pressing a button, giving them more autonomy and removing the need for continuous step-by-step instructions from the experimenter. This high degree of automation significantly reduced the burden on experimenters and minimized errors from manual intervention. Another advantage was the possibility to test multiple participants in parallel without any risk of cross-exposure to stray light from other participants taking part. All these factors vastly accelerated the entire data-gathering process.

Experimental design and stimulus variety

The programmability of the displays also allowed us to present more complex stimuli as part of a larger study, such as temporally modulated light with different inputs to each eye. Other possibilities would include presenting spatially modulated stimuli, or even performing more complex visual tasks such as visual attention tests. Another added benefit was the possibility to randomize the order of the conditions while experimenters remained blinded to which stimulus was being presented each session.

Participant comfort

The use of HMDs also gave participants greater freedom of movement while being exposed to the stimulus, increasing comfort. Traditional highly controlled lab-based setups such as the ganzfeld dome stimulator, often require participants to remain in a fixed position in front of a light source throughout the duration of the intervention. This constraint is not present HMDs, allowing participants to sit more comfortably and to change their posture, while maintaining the same visual stimulation. This is particularly key for circadian experiments due to the long duration of experimental protocols.

Implications for circadian research

Our findings confirm that VR HMD-based light delivery can effectively induce acute melatonin suppression, a key endpoint in circadian and sleep research. By precisely regulating light intensity and spectral composition, researchers can better isolate the physiological effects of specific lighting conditions. This method also opens the door to investigating advanced protocols that incorporate flicker or other temporally or spatially varying light stimuli, potentially providing insight into how dynamic lighting patterns influence circadian phase and other physiological rhythms.

Standardized stimuli across multi-centric studies and projects

The proposed approach facilitates the standardization of stimuli across different study sites. It is generally non-trivial to equate stimulus conditions across different laboratory settings and technological specifications. For example, when using room illumination as a stimulus, the exact dimensions of a room and light-related parameters such as the wall reflectance will modify the corneal light exposure received by a participant. Even when using integrating spheres with light sources, there is typically no guarantee that the stimuli received by an observer will be the same across two different setups. Using standard consumer hardware to display stimuli solves this issue, thereby opening the door to collect harmonized data linked by a common stimulus display system across multiple sites.

Higher efficiency in in-laboratory data collection

A key feature of experiments targeting neuroendocrine and circadian physiology is that they are costly in terms of time and resources. Using VR displays makes the collection of data a lot more efficient, due to the possibilities for automation and the ability to run multiple participants at a time.

Scalability for remote and at-home data collection

An especially promising avenue is the use of VR HMDs for data collection outside the laboratory. When combined with user-friendly calibration protocols and remote monitoring (e.g., via embedded eye tracking or integrated sensors), participants could complete lighting studies and interventions at home with scientifically validated stimuli. This approach would facilitate large-scale studies and longitudinal designs, expanding research capacity to populations that may otherwise be unable or unwilling to travel to specialized research facilities.

Limitations and future directions

Although the present study demonstrates the feasibility of using VR displays for controlled light delivery, certain limitations must be addressed. First, the inbuilt strobing at 90 Hz found in the devices used, which could not be disabled, poses challenges for high-frequency flicker stimuli above approximately 20 Hz. Other headset models might not have this feature or offer options for disabling it for improved temporal precision.

We also found minor discrepancies in the gamma function of the different colour primaries, which could mean that using these displays is less suitable for studies where high accuracy of the individual display primaries is needed, such as for the silent substitution method. However, this could be addressed with more extensive light measurements and a more complex calibration procedure.

There was a subset of participants for which the light stimulation did not produce an acute suppression effect on melatonin. This could be addressed by using a higher intensity, since the stimulus used in this study was only 50% of the maximum brightness of the HMD. Potentially, the use of pupil dilation could also increase the suppression effect by increasing retinal illuminance (Gaddy et al., 1993), although within our sample this increase was not statistically significant.

Beyond these limitations found in our study, other potential limitations of HMDs require careful consideration. While the HMD used in this study allowed for individual adjustment of the interpupillary distance, and this was a required step as part of the inbuilt calibration procedure, this feature is not present in all commercial devices. This could introduce discrepancies in the visual stimulus received by different participants. In general, establishing standardized fitting protocols and using robust calibration procedures, together with providing clear adjustment instructions to participants, are necessary steps in order to ensure the precision of this methodology. Another characteristic of most HMDs is that they require straps to be fitted around the head of participants, making them rather unsuitable for studies that use

electroencephalogram, facial electromyography, or other similar techniques that require free access to the face or the head.

Finally, future research should assess a larger variety of commercially available devices. Different VR headsets have varying hardware specifications and different level of control over the image being presented. Furthermore, different models would fundamentally differ in their level of brightness, resolution and field of view. Future work should assess which devices among those available are suitable for use in research, particularly for multi-site or large-scale studies.

Conclusion

Our results demonstrate the promising potential of commercial head-mounted displays as a highly controlled, scalable, and standardized tool for investigating non-visual effects of light such as melatonin suppression. By combining meticulous temporal and spectral calibration, and the capacity for multi-site and remote data collection, VR displays can significantly enhance the methodological capabilities of chronobiology and sleep research.

References

Adhanom, I. B., MacNeilage, P., & Folmer, E. (2023). Eye Tracking in Virtual Reality: A Broad Review of Applications and Challenges. *Virtual Reality*, *27*(2), 1481–1505. https://doi.org/10.1007/s10055-022-00738-z

Barrionuevo, P. A., Preciado, O. U., Sandoval Salinas, M. L., & Issolio, L. A. (2022). Optical stimulation systems for studying human vision. *Progress in Brain Research*, *273*(1), 13–36. https://doi.org/10.1016/bs.pbr.2022.04.003

Berson, D. M., Dunn, F. A., & Takao, M. (2002). Phototransduction by retinal ganglion cells that set the circadian clock. *Science (New York, N.Y.)*, *295*(5557), 1070–1073. https://doi.org/10.1126/science.1067262

Blume, C., Cajochen, C., Schöllhorn, I., Slawik, H. C., & Spitschan, M. (2024). Effects of calibrated blue–yellow changes in light on the human circadian clock. *Nature Human Behaviour*, *8*(3), 590–605.

Blume, C., Garbazza, C., & Spitschan, M. (2019). Effects of light on human circadian rhythms, sleep and mood. *Somnologie*, *23*(3), 147–156. https://doi.org/10.1007/s11818-019-00215-x

Brainard, G. C., Hanifin, J. P., Greeson, J. M., Byrne, B., Glickman, G., Gerner, E., & Rollag, M. D. (2001). Action spectrum for melatonin regulation in humans: Evidence for a novel circadian photoreceptor. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *21*(16), 6405–6412. https://doi.org/10.1523/JNEUROSCI.21-16-06405.2001

Brown, T. M. (2020). Melanopic illuminance defines the magnitude of human circadian light responses under a wide range of conditions. *Journal of Pineal Research*, *69*(1), e12655. https://doi.org/10.1111/jpi.12655

Brown, T. M., Brainard, G. C., Cajochen, C., Czeisler, C. A., Hanifin, J. P., Lockley, S. W., Lucas, R. J., Münch, M., O'Hagan, J. B., Peirson, S. N., Price, L. L. A., Roenneberg, T., Schlangen, L. J. M., Skene, D. J., Spitschan, M., Vetter, C., Zee, P. C., & Wright, K. P. (2022). Recommendations for daytime, evening, and nighttime indoor light exposure to best support physiology, sleep, and wakefulness in healthy adults. *PLoS Biology*, *20*(3), e3001571. https://doi.org/10.1371/journal.pbio.3001571

Chia, Z. K., Kong, A. W., Turner, M. L., Saifee, M., Damato, B. E., Backus, B. T., Blaha, J. J., Schuman, J. S., Deiner, M. S., & Ou, Y. (2023). Assessment of Remote Training, At-Home Testing, and Test-Retest Variability of a Novel Test for Clustered Virtual Reality Perimetry. *Ophthalmology Glaucoma*, S2589419623001606. https://doi.org/10.1016/j.ogla.2023.08.006

Creem-Regehr, S., Kelly, J., Bodenheimer, B., & Stefanucci, J. (2023). *Virtual Reality as a Tool to Understand Spatial Navigation*. https://doi.org/10.31234/osf.io/djhmw

Creem-Regehr, S., Stefanucci, J. K., & Bodenheimer, B. (2022). Perceiving distance in virtual reality: Theoretical insights from contemporary technologies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1869), 20210456. https://doi.org/10.1098/rstb.2021.0456

Cwierz, H., Díaz-Barrancas, F., Llinás, J. G., & Pardo, P. J. (2021). On the Validity of Virtual Reality Applications for Professional Use: A Case Study on Color Vision Research and Diagnosis. *IEEE Access*, *9*, 138215–138224. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3118438

Dacey, D. M., Liao, H.-W., Peterson, B. B., Robinson, F. R., Smith, V. C., Pokorny, J., Yau, K.-W., & Gamlin, P. D. (2005). Melanopsin-expressing ganglion cells in primate retina signal colour and irradiance and project to the LGN. *Nature*, *433*(7027), 749–754. https://doi.org/10.1038/nature03387

De Gelder, B., Kätsyri, J., & De Borst, A. W. (2018). Virtual reality and the new psychophysics. *British Journal of Psychology*, *109*(3), 421–426. https://doi.org/10.1111/bjop.12308

Do, M. T. H. (2019). Melanopsin and the Intrinsically Photosensitive Retinal Ganglion Cells: Biophysics to Behavior. *Neuron*, *104*(2), 205–226. https://doi.org/10.1016/j.neuron.2019.07.016

Do, M. T. H., Kang, S. H., Xue, T., Zhong, H., Liao, H.-W., Bergles, D. E., & Yau, K.-W. (2009). Melanopsin Signalling: Low Pigment Density, Large Single-Photon Response, and High-Efficiency Transmission. *Biophysical Journal*, *96*(3), 200a. https://doi.org/10.1016/j.bpj.2008.12.1074

Gaddy, J. R., Rollag, M. D., & Brainard, G. C. (1993). Pupil size regulation of threshold of light-induced melatonin suppression. *The Journal of Clinical Endocrinology & Metabolism*, *77*(5), 1398–1401.

Hattar, S., Liao, H. W., Takao, M., Berson, D. M., & Yau, K. W. (2002). Melanopsin-containing retinal ganglion cells: Architecture, projections, and intrinsic photosensitivity. *Science (New York, N.Y.)*, *295*(5557), 1065–1070. https://doi.org/10.1126/science.1069609

Higuchi, S., Ishibashi, K., Aritake, S., Enomoto, M., Hida, A., Tamura, M., Kozaki, T., Motohashi, Y., & Mishima, K. (2008). Inter-individual difference in pupil size correlates to suppression of melatonin by exposure to light. *Neuroscience Letters*, *440*(1), 23–26. https://doi.org/10.1016/j.neulet.2008.05.037

Hu, G. Y., Prasad, J., Chen, D. K., Alcantara-Castillo, J. C., Patel, V. N., & Al-Aswad, L. A. (2023). Home Monitoring of Glaucoma Using a Home Tonometer and a Novel Virtual Reality Visual Field Device: Acceptability and Feasibility. *Ophthalmology Glaucoma*, *6*(2), 121–128. https://doi.org/10.1016/j.ogla.2022.05.001

Lepori, F., Maiello, G., Dehn, K., Hartmann, F., Chessa, M., Rothkopf, C. A., & Fleming, R. W. (2023). A method to align real and virtual objects for mixed reality investigations of visually guide grasping. *Journal of Vision*, *23*(9), 4797. https://doi.org/10.1167/jov.23.9.4797

Lucas, R. J., Peirson, S. N., Berson, D. M., Brown, T. M., Cooper, H. M., Czeisler, C. A., Figueiro, M. G., Gamlin, P. D., Lockley, S. W., O'Hagan, J. B., Price, L. L. A., Provencio, I., Skene, D. J., & Brainard, G. C. (2014). Measuring and using light in the melanopsin age. *Trends in Neurosciences*, *37*(1), 1–9. https://doi.org/10.1016/j.tins.2013.10.004

Mees, L., Upadhyaya, S., Kumar, P., Kotawala, S., Haran, S., Rajasekar, S., Friedman, D. S., & Venkatesh, R. (2020). Validation of a Head-mounted Virtual Reality Visual Field Screening Device. *Journal of Glaucoma*, *29*(2), 86–91. https://doi.org/10.1097/IJG.0000000000001415

Neugebauer, A., Stingl, K., Ivanov, I., & Wahl, S. (2021). Influence of Systematic Gaze Patterns in Navigation and Search Tasks with Simulated Retinitis Pigmentosa. *Brain Sciences*, *11*(2), 223. https://doi.org/10.3390/brainsci11020223

Prayag, A. S., Najjar, R. P., & Gronfier, C. (2019). Melatonin suppression is exquisitely sensitive to light and primarily driven by melanopsin in humans. *Journal of Pineal Research*, *66*(4), e12562. https://doi.org/10.1111/jpi.12562

Schöllhorn, I., Stefani, O., Lucas, R. J., Spitschan, M., Epple, C., & Cajochen, C. (2024). The Impact of Pupil Constriction on the Relationship Between Melanopic EDI and Melatonin Suppression in Young Adult Males. *Journal of Biological Rhythms*, *39*(3), 282–294. https://doi.org/10.1177/07487304241226466

Schöllhorn, I., Stefani, O., Lucas, R. J., Spitschan, M., Slawik, H. C., & Cajochen, C. (2023). Melanopic irradiance defines the impact of evening display light on sleep latency, melatonin and alertness. *Communications Biology*, *6*(1), 228. https://doi.org/10.1038/s42003-023-04598-4

Spitschan, M. (2019). Melanopsin contributions to non-visual and visual function. *Current Opinion in Behavioral Sciences*, *30*, 67–72. https://doi.org/10.1016/j.cobeha.2019.06.004

Spitschan, M. (2021). Time-Varying Light Exposure in Chronobiology and Sleep Research Experiments. *Frontiers in Neurology*, *12*. https://www.frontiersin.org/articles/10.3389/fneur.2021.654158

Spitschan, M., Kervezee, L., Lok, R., McGlashan, E., Najjar, R. P., & ENLIGHT Consortium. (2023). ENLIGHT: A consensus checklist for reporting laboratory-based studies on the non-visual effects of light in humans. *EBioMedicine*, *98*, 104889. https://doi.org/10.1016/j.ebiom.2023.104889

Tatiyosyan, S. A., Rifai, K., & Wahl, S. (2020). Standalone cooperation-free OKN-based low vision contrast sensitivity estimation in VR - a pilot study. *Restorative Neurology and Neuroscience*, *38*(2), 119–129. https://doi.org/10.3233/RNN-190937

Thapan, K., Arendt, J., & Skene, D. J. (2001). An action spectrum for melatonin suppression: Evidence for a novel non-rod, non-cone photoreceptor system in humans. *The Journal of Physiology*, *535*(Pt 1), 261–267. https://doi.org/10.1111/j.1469-7793.2001.t01-1-00261.x

Thibos, L. N., Lopez-Gil, N., & Bradley, A. (2018). What is a troland? *Journal of the Optical Society of America A*, *35*(5), 813. https://doi.org/10.1364/josaa.35.000813

Tsapakis, S., Papaconstantinou, D., Diagourtas, A., Kandarakis, S., Droutsas, K., Andreanos, K., & Brouzas, D. (2018). Home-based visual field test for glaucoma screening comparison with Humphrey perimeter. *Clinical Ophthalmology*, *Volume 12*, 2597–2606. https://doi.org/10.2147/OPTH.S187832

Wiley, E., Khattab, S., & Tang, A. (2022). Examining the effect of virtual reality therapy on cognition post-stroke: A systematic review and meta-analysis. *Disability and Rehabilitation: Assistive Technology*, *17*(1), 50–60. https://doi.org/10.1080/17483107.2020.1755376

Supplementary Materials

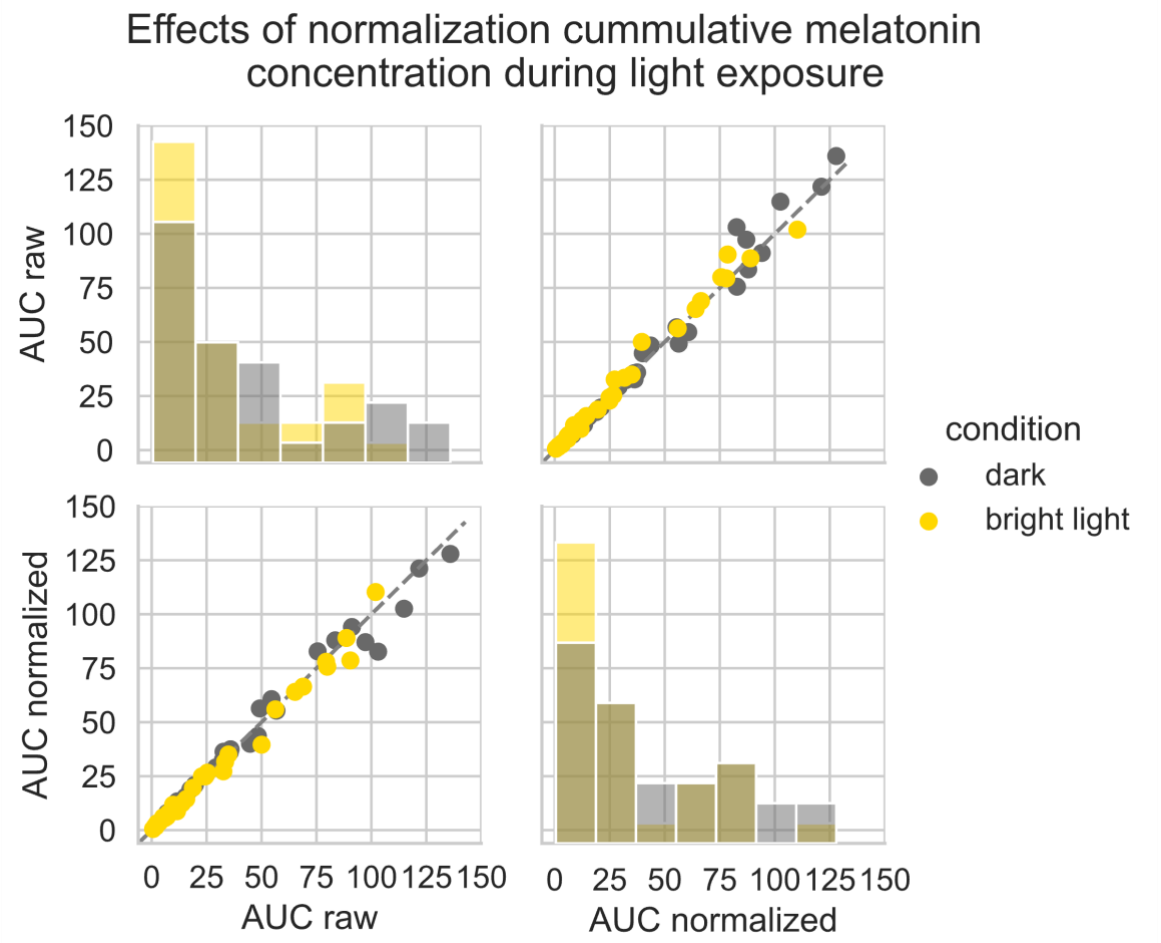

Figure S1. Comparison of the effects of the normalization procedure implemented (see Methods section) on the AUCs or cumulative melatonin concentrations during light exposure.

Author contributions

Conceptualization, M.F.A. and M.S.; Methodology, M.F.A. and M.S.; Software, M.F.A.; Validation, M.F.A. and M.S.; Formal Analysis, M.F.A.; Investigation, M.F.A. and M.S.; Resources, M.S.; Data Curation, M.F.A; Writing – Original Draft Preparation, M.F.A.; Writing

– Review & Editing, M.F.A. and M.S.; Visualization, M.F.A.; Supervision, M.S.; Project Administration, M.F.A. and M.S.; Funding Acquisition, M.S.

Funding

This research was funded by the Max Planck Society in the form of a free-floating Max Planck Research Group (M.S.).

Ethics statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the Technical University of Munich (2022-439-S-SR).

Acknowledgments

We thank Jennifer Reif, Nik Novik and Sing Yik Chan for their contributions to this project. Furthermore, we thank the staff at the mechanical and electronics workshop of the Max Planck Institute for Biological Cybernetics for their technical support.

Conflicts of Interest

The authors declare no conflict of interest.