Different rules for binocular combination of luminance flicker in cortical and subcortical pathways

Federico G. Segala, Aurelio Bruno, Myat T. Aung, Alex R. Wade & Daniel H. Baker

Department of Psychology, University of York, UK

1 Abstract

How does the human brain combine information across the eyes? It has been known for many years that cortical normalisation mechanisms implement 'ocularity invariance': equalising neural responses to spatial patterns presented either monocularly or binocularly. Here we used a novel combination of electrophysiology, psychophysics, pupillometry and computational modelling to ask whether this invariance also holds for flickering luminance stimuli with no spatial contrast. We find dramatic violations of ocularity invariance for these stimuli, both in cortex and also in the subcortical pathways that govern pupil diameter. Specifically, we find substantial binocular facilitation in both pathways with the effect being strongest in cortex. Near-linear binocular additivity (instead of ocularity invariance) was also found using a perceptual luminance matching task. Ocularity invariance is therefore not a ubiquitous feature of visual processing, and the brain appears to repurpose a generic normalisation algorithm for different visual functions by adjusting the amount of interocular suppression.

2 Introduction

The brain must combine information across multiple sensory inputs to derive a coherent percept of the external world. This involves a process of signal combination both within (Baker and Wade, 2017) and between (Ernst and Banks, 2002) the senses. Binocular vision is a useful test case for signal combination, as the inputs to the two eyes overlap substantially (in species with forward-facing eyes), and the neural locus is well-established (Hubel and Wiesel, 1962). Much of our knowledge about binocular combination derives from studies on the contrast response of the 'canonical' visual pathway, in which signals pass from the eyes to primary visual cortex (V1), via the lateral geniculate nucleus (LGN) (Purves et al., 2008). However, signals are also combined across the eyes in the network of subcortical nuclei that govern pupil diameter in response to absolute light levels (McDougal and Gamlin, 2008) and much less is known about the computations that operate in these subcortical pathways. In addition, our primary purpose here is to investigate the computations governing signal combination in these two anatomically distinct pathways in response to luminance changes.

For pattern vision, binocular presentation confers greater sensitivity to low contrast targets than monocular presentation. This is known as binocular summation, with summation ratios (the relative improvement under binocular presentation) at detection threshold lying between $\sqrt{2}$ and 2 (Baker et al., 2018; Campbell and Green, 1965). This advantage is lost at high stimulus intensities, where both psychophysical performance (contrast discrimination thresholds) (Legge, 1984; Meese et al., 2006) and neural activity (Baker and Wade, 2017; Moradi and Heeger, 2009) are approximately equal for monocular and binocular presentation. Contemporary models of binocular vision (Ding and Sperling, 2006; Meese et al., 2006) advocate a process of interocular suppression that normalizes the two eyes inputs at high contrasts and negates the binocular

advantage. This is consistent with our everyday experience of 'ocularity invariance' (Baker et al., 2007): perceived contrast does not change when one eye is opened and closed.

The pupillary light reflex is an automatic constriction of the iris sphincter muscles in response to increases in light levels, which causes the pupil to shrink (McDougal and Gamlin, 2008). There is a clear binocular component to this reflex, as stimulation of one eye still causes constriction of the other eye's pupil (termed the consensual response; Wyatt and Musselman (1981)). Importantly, the neuroanatomical pathway involved completely bypasses the canonical cortical pathway (retina to V1), instead involving a network of subcortical nuclei, including the Pretectal Olivary nucleus, Superior Cervical ganglion, and Edinger-Westphal nucleus (Angée et al., 2021; Mathôt, 2018; McDougal and Gamlin, 2008; Wang and Munoz, 2015). To account for the consensual response, these brain regions must combine information from the left and right eyes (Doesschate and Alpern, 1967), yet the computation that achieves this is achieved is unclear.

To address this shortcoming, we designed an experiment that allowed us to simultaneously record electrophysiological and pupillometric responses to monocular and binocular stimuli. This novel paradigm allowed us to probe both cortical (using EEG) and subcortical (using a binocular eyetracker) pathways simultaneously in response to flickering light, and make quantitative comparisons between them. Periodic flicker entrains both neural (Norcia et al., 2015) and pupil (Spitschan et al., 2014) responses at the flicker frequency, enabling precise estimation of response amplitudes in the Fourier domain. We followed up our main experiment with additional exploration of the effect of stimulus frequency, and a psychophysical matching experiment measuring perceived flicker intensity (i.e. temporal contrast). The results are interpreted using a hierarchical Bayesian computational model of binocular vision, and reveal that subcortical pathways implement stronger interocular suppression than the canonical cortical pathway.

3 Methods

3.1 Participants

Thirty (20 females), twelve (7 females) and ten (3 females) adult participants, whose ages ranged from 18 to 45, were recruited for Experiments 1, 2 and 3 respectively. All participants had normal or corrected to normal binocular vision, and gave written informed consent. Our procedures were approved by the Ethics Committee of the Department of Psychology at the University of York (identification number 792).

3.2 Apparatus & stimuli

The stimuli were two discs of achromatic flickering light with a diameter of 3.74 degrees, presented on a black background. The same stimuli were used for all three experiments. Four dark red lines were added around both discs to help with their fusion into one binocular disc (see insert in Figure 1b for an example of the fused stimulus). The discs were viewed through a four-mirror stereoscope, which used front silvered mirrors to avoid internal reflections, and meant that participants saw a single fused disc. The use of a stereoscope allowed us to modulate the stimuli in three different ocular configurations: monocular, binocular, and dichoptic. Note that during monocular presentation of flicker, the unstimulated eye still saw the static (non-flickering) disc of mean luminance.

All stimuli had a mean luminance of 42 cd/m^2 on an Iiyama VisionMaster™ Pro 510 display (800 x 600 pixels, 60 Hz refresh rate), which was gamma corrected using a Minolta LS-110 photometer (Minolta Camera Co. Ltd., Japan). For experiments 1 and 2, the stimuli were presented using Psychopy (v3.0.7). For experiment 3, the stimuli were presented using Psychopy (v2022.1.1).

EEG data were collected for Experiments 1 and 2 using a 64-electrode ANT WaveGuard cap and the signals were recorded at 1 kHz using the ASA software (ANT Neuro, Netherlands). Pupillometry data were collected for Experiment 1 using a binocular Pupil Core eye-tracker (Pupil Labs GmbH, Berlin, Germany; Kassner et al. (2014)) running at 120 Hz, and the signals were recorded with the Pupil Capture software.

3.3 Procedure

Before each experiment, participants calibrated the stereoscope by adjusting the angle of the mirrors. This was done so that they would perceive the two discs as one fused disc when looking at the screen through the stereoscope.

3.3.1 Experiment 1: simultaneous EEG and pupillometry

The experiment was conducted in a windowless room, in which the only light source was the monitor. The participants sat at 99 cm from the monitor and the total optical viewing distance (through the stereoscope) was 107 cm. The experiment was carried out in a single session lasting 45 minutes in total, divided in three blocks of 15 minutes each. In each block, there were 60 trials lasting 15 seconds each (12s of stimulus presentation, with an interstimulus interval of 3s). The participants were given no task other than to look at the fixation cross in the middle of the disc while trying to minimise their blinking during the presentation period.

We included six distinct ocular conditions, each at five temporal contrast levels (combined factorially) relative to the mean luminance: 6, 12, 24, 48 and 96%. Contrast was defined as temporal Michelson contrast; the difference between maximum and minimum luminances, scaled by the mean and expressed as a percentage. In the first three conditions, the discs flickered at 2 Hz, in either a monocular, binocular, or dichoptic arrangement. In the dichoptic condition the non-target eye saw a fixed contrast of 48%. In the remaining three conditions (termed the cross-frequency conditions) one eye's disc flickered at 1.6Hz, and the other eye's disc flickered at 2Hz. We also tested monocular (one eye sees 1.6Hz flicker, the other sees mean luminance), binocular (one eye sees each frequency at the target contrast) and dichoptic (target stimulus flickering at 2Hz, mask contrast of 48% at 1.6Hz in the other eye) arrangements. The rationale for flickering both eyes at 2Hz is that we can then measure summation behaviour between the eyes in the pupil and EEG response at 2Hz. The rationale for flickering the eyes at different frequencies is that this permits measurement of suppression between the eyes (i.e. the reduction in the 2Hz response when a 1.6Hz mask component is added to the other eye). We counterbalanced presentation of the target stimulus across the left and right eyes.

3.3.2 Experiment 2: EEG responses across temporal frequency

This experiment used the same equipment set up as Experiment 1, except that the eye tracker was not used. Unlike the first experiment, only one contrast level was used (96%) and the discs were set to flicker at five different frequencies (2, 4, 8, 16 and 30 Hz). Only two ocular configurations, monocular and binocular, were included, with the latter having both discs flickering at the same frequency. The experiment was carried out in one session lasting 25 minutes in total, divided into five blocks of 5 minutes each. In each block, there were 20 trials in total with the same timing as for Experiment 1.

3.3.3 Experiment 3: temporal contrast matching

The experiment was conducted in a darkened room with a blacked-out window. The display equipment (monitor and stereoscope) were the same as for the two previous experiments, but no EEG or pupillometry data were collected. A two-interval contrast matching procedure was used to collect data. In one interval, participants were presented with a standard fused disc that flickered at a set contrast level (either 24 or 48%), which was selected by the experimenter at the beginning of each block. In the other interval, a target disc flickering at varying contrast levels was displayed. The contrast level of the target was controlled by a 1-up, 1-down staircase moving in logarithmic (dB) steps of contrast. The ratio of flicker amplitudes in the left and right eyes was varied across blocks and was set to be 0, 0.25, 0.5, 0.75 or 1 (9 distinct conditions). The standard and target discs were displayed for 1 second each, with an interstimulus interval of 0.5 seconds. After both discs had appeared on screen, the participants had to indicate which interval they perceived as having the more intense flicker. The intervals were randomly ordered, and all discs flickered at a frequency of 2 Hz (two cycles in sine phase).

Due to its long duration (approximately 3 hours in total), the participants completed the experiment across multiple sessions initiated at their own convenience. The experiment was divided into 54 blocks (3 repetitions $\times$ 2 standard contrasts $\times$ 9 target ratios), which lasted on average 3 minutes each, depending on the response speed of the participant. In each block, there were a total of 50 trials. No auditory feedback was given for this subjective task.

3.4 Data analysis

EEG data were converted from the ANT-EEProbe format to a compressed csv text file using a custom Matlab script and components of the EEGlab toolbox (Delorme and Makeig, 2004). The data for each participant were then loaded into R for analysis, where a ten-second waveform for each trial at each electrode was extracted (omitting the first two seconds). The Fourier transform of each waveform was calculated, and the complex spectrum stored in a matrix. All repetitions of each condition were then averaged for each electrode. They were then averaged across four occipital electrodes (*POz*, *Oz*, *O1*, *O2*), to obtain individual results. Finally, these were averaged across participants to obtain the group results. All averaging was performed in the complex domain and therefore retained the phase information (i.e. coherent averaging), and at each stage we excluded data points with a Mahalanobis distance exceeding $D = 3$ from the complex-valued mean (see Baker, 2021). For statistical comparisons of complex-valued data, we use the $ANOVA^2_{circ}$ statistic described by Baker (2021). This is a multivariate extension of ANOVA that assumes equal variance of the real and imaginary Fourier components, or equivalently, an extension of the T^2_{circ} statistic of Victor and Mast (1991) that can compare more than two conditions.

A similar analysis pipeline was adopted for the pupillometry data. The data were converted from mp4 videos to a csv text file using the Pupil Player software (Kassner et al., 2014), which estimated pupil diameter for each eye on each frame using a 3D model of the eyeball. The individual data were then loaded into R for analysis, where again a ten-second waveform for each trial in each eye was extracted (excluding the first two seconds after stimulus onset). We interpolated across any dropped or missing frames to ensure regular and continuous sampling over time. The Fourier transform was calculated for each waveform, and all repetitions of each condition were pooled across eye and then averaged. Finally, data were averaged across all participants to obtain the group results. Again, we used coherent averaging, and excluded outlying data points in the same way as for the EEG data. Note that previous pupillometry studies using luminance flicker have tended to fit a single sine-wave at the fundamental frequency, rather than using Fourier analysis (e.g. Spitschan et al., 2014). The Fourier approach is more robust to noise at other frequencies and has been used in some previous studies (see Barrionuevo et al., 2014; Barrionuevo and Cao, 2016). Additionally, it makes the pupillometry analysis consistent with standard practice in steady-state EEG analysis (e.g. Figueira et al., 2022).

To analyse the matching data, we pooled the trial responses across all repetitions of a given condition for each participant. We then fitted a cumulative normal psychometric function to estimate the point of subjective equality at the 50% level. Thresholds were averaged across participants in logarithmic (dB) units.

For all experiments, we used a bootstrapping procedure with 1000 iterations to estimate standard errors across participants. All analysis and figure construction was conducted using a single R-script, available online, making this study fully computationally reproducible.

3.5 Computational model and parameter estimation

To describe our data, we chose a model of binocular contrast gain control with the same general form as the first stage of the model proposed by Meese et al. (2006). The second gain control stage was omitted (consistent with Baker and Wade, 2017) to simplify the model and reduce the number of free parameters. The response of the left eye's channel is given by:

$$Resp_L = \frac{L^2}{Z + L + wR}, \tag{1}$$

with an equivalent expression for the right eye:

$$Resp_R = \frac{R^2}{Z + R + wL}. \tag{2}$$

In both equations, L and R are the contrast signals from the left and right eyes, Z is a saturation constant that shifts the contrast-response function laterally, and w is the weight of suppression from the other eye.

The responses from the two eyes are then summed binocularly:

$$Resp_B = R_{max}(Resp_L + Resp_R) + n, \tag{3}$$

where n is a noise parameter, and R_{max} scales the overall response amplitude. The R_{max} parameter was omitted when modelling the contrast matching data, as it has no effect in this paradigm.

Despite being derived from the model proposed by Meese et al. (2006), the simplifications applied to this architecture make it very similar to other models (e.g. Ding and Sperling, 2006; Doesschate and Alpern, 1967; Legge, 1984; Schrödinger, 1926). In particular we fixed the numerator exponent at 2 in our model, because otherwise this value tends to trade off with the weight of interocular suppression (see Baker et al., 2012; Kingdom and Libenson, 2015). Our key parameter of interest is the weight of interocular suppression. Large values around $w = 1$ result in a very small or nonexistent binocular advantage at suprathreshold contrasts, consistent with previous work using grating stimuli (Baker and Wade, 2017). Low values around $w = 0$ produce substantial, near-linear binocular facilitation (Baker et al., 2020).

We implemented the model within a Bayesian framework using the Stan software (Carpenter et al., 2017). This allowed us to estimate group-level posterior parameter distributions for the weight of interocular suppression, w, and the other free model parameters R_{max}, Z and n. The prior distributions for all parameters were Gaussian, with means and standard deviations of 1 and 0.5 for w and R_{max}, and 5 and 2 for Z and n, with these values chosen based on previous literature (Baker et al., 2012; Meese et al., 2006). We sampled from a Student's t-distribution for the amplitudes in the pupillometry and EEG experiments, and from a Bernoulli distribution for the single trial matching data. The models were fit using the individual data across all participants, independently for each data set. Because we used coherent averaging across participants, the group average amplitudes are shifted vertically relative to the model predictions, which are based on hierarchical fits to the individual participant amplitudes (put another way, the model does not implement coherent averaging). However it is clear that the character of the model in all cases gives a good representation of the data, as can be seen in Figure 6. We took posterior samples at over a million steps for each data set, using a computer cluster, and retained 10% of samples for plotting.

3.6 Preregistration, data and code availability

We initially preregistered our main hypotheses and analysis intentions for the first experiment. We then conducted a pilot study with N=12 participants, before making some minor changes to the stimulus (we added dim red lines to aid binocular fusion). We then ran the main experiment, followed by two additional experiments that were not preregistered. The preregistration document, raw data files, and experimental and analysis code, are available on the project repository: https://doi.org/10.17605/OSF.IO/TBEMA.

4 Results

4.1 Experiment 1

The pupillometry results are summarised in Figure 1. The group average waveform for binocular presentation is shown in Figure 1a. There is a substantial pupil constriction at stimulus onset, followed by visible

oscillations at the flicker frequency (2Hz, see waveform at foot). The average Fourier spectrum is displayed in Figure 1b, and shows a clear spike at 2 Hz, but no evidence of a second harmonic response at 4Hz. These results demonstrate that our paradigm can evoke measurable steady-state pupil responses at 2Hz.

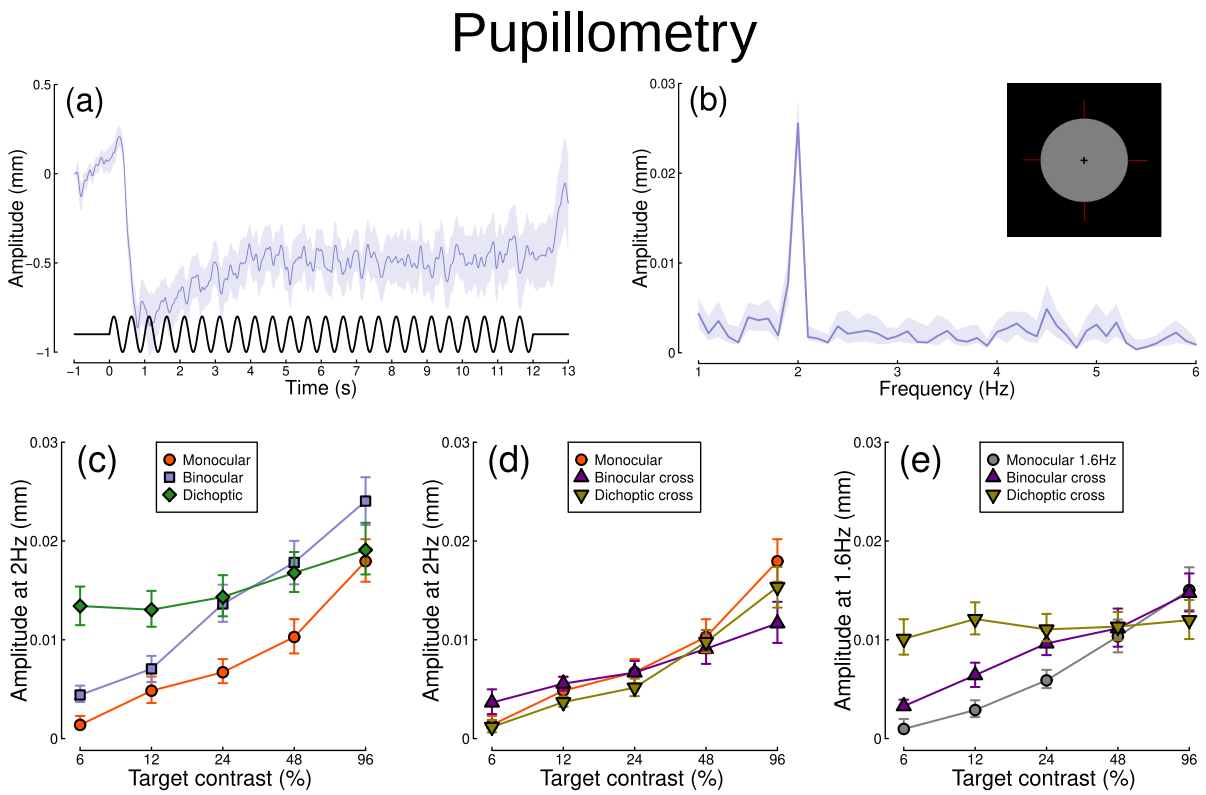

Figure 1: Summary of pupillometry results for N=30 participants. Panel (a) shows a group average waveform for binocular presentation (low pass filtered at 5Hz), with the driving signal plotted at the foot. Panel (b) shows the average Fourier spectrum, with an inset image illustrating the stimulus appearance (upper right). Panels (c,d) show contrast response functions at 2Hz for different conditions. Panel (e) shows contrast response functions at 1.6Hz for three conditions. Shaded regions and error bars indicate bootstrapped standard errors.

Figure 1c shows contrast response functions in response to stimuli flickering only at 2Hz. Response amplitudes increased monotonically with target contrast, confirming that our paradigm is suitable for measuring contrast-dependent differences in response (to our knowledge this is the first time this has been demonstrated). The amplitude of the binocular condition (blue squares) is consistently greater than that of the monocular condition (red circles) across all target contrasts. A 2×5 repeated measures $ANOVA^2_{circ}$ (Baker, 2021) comparing these conditions revealed a significant main effect of target contrast (F(8,580) = 16.79, $p < 0.001$), a significant effect of condition (F(2,580) = 11.04, $p < 0.001$), and a significant interaction (F(8,580) = 56.25, $p < 0.001$). The dichoptic condition begins at a much higher amplitude, owing to binocular combination of the target and high (48%) contrast mask, and then increases slightly with increasing target contrast (main effect of target contrast: F(8,232) = 3.03, $p < 0.003$).

In Figure 1d, we plot responses to monocular target stimuli flickering at 2Hz, when the other eye viewed stimuli flickering at 1.6Hz (the red monocular-only data are replotted from Figure 1c for comparison). When the 1.6Hz component had the same contrast as the target (the binocular cross condition, shown in purple) responses were facilitated slightly at low contrasts, and suppressed at the highest target contrasts (interaction between contrast and condition: F(8,580) = 52.94, $p < 0.001$). When the 1.6Hz component had a fixed contrast of 48% (the dichoptic cross condition, shown in yellow), responses were suppressed slightly

across the contrast range (interaction between contrast and condition: F(8,580) = 62.05, $p < 0.001$).

Figure 1e shows responses at 1.6Hz, for the same conditions, as well as for a condition in which a monocular stimulus flickered at 1.6Hz (grey circles). Surprisingly there again appears to be a slight facilitation effect in the binocular cross condition, particularly at lower contrasts. The dichoptic cross condition does not show clear modulation with target contrast.

Figure 2 shows equivalent results, measured contemporaneously using EEG. Figure 2a shows the group average waveform for binocular presentation, and Figure 2b shows the Fourier spectrum for binocular presentation, both averaged across four posterior electrodes (*Oz*, *POz*, *O1* and *O2*, marked on the inset scalp plots). Unlike for the pupillometry data, there are clear responses at both the first harmonic frequency (2Hz), and also the second harmonic frequency (4Hz). We therefore calculated contrast response functions at both first and second harmonic frequencies.

When stimuli in both eyes flicker at 2Hz, the binocular responses at the first (Figure 2c) and second (Figure 2f) harmonics are substantially greater than the monocular responses, particularly at high contrasts. Analysis of variance on the complex values ($ANOVA^2_{circ}$) revealed a main effect of contrast (F(8,580) = 4.38, $p < 0.001$) and an interaction effect (F(8,580) = 61.58, $p < 0.001$), but no effect of condition ($p = 0.13$) at the first harmonic, with a similar pattern of results obtained at the second harmonic. For the cross-frequency conditions (Figure 2d,g), there was no appreciable effect of adding a 1.6Hz component on the response at 2Hz or 4Hz (no effect of condition, and no interaction). Similarly, there were no clear interocular interactions between frequencies in the responses at 1.6Hz (Figure 2e) and 3.2Hz (Figure 2h). This pattern of results suggests that processing of temporal luminance modulations happens in a more linear way in visual cortex (indexed by EEG), compared with subcortical pathways (indexed by pupillometry), and shows no evidence of interocular suppression.

Finally, we calculated the ratio of binocular to monocular responses across the three data types from Experiment 1. Figure 3 shows that these ratios are approximately $\sqrt{2}$ across the low-to-intermediate contrast range for all three data types. At higher contrasts, we see ratios of 2 or higher for the EEG data, but much weaker ratios near 1 for the pupillometry data. Note that the ratios here are calculated on a per-participant basis and then averaged, rather than being the ratios of the average values shown in Figures 1 and 2. A 3×5 repeated measures ANOVA on the logarithmic (dB) ratios found a main effect of contrast (F(3.08,89.28) = 4.53, $p < 0.002$), no effect of data modality (F(2,58) = 0.75, $p = 0.48$), but a highly significant interaction (F(5.54,160.67) = 3.84, $p < 3 \times 10^{-4}$).

4.2 Experiment 2

The strong binocular facilitation and weak interocular suppression in the EEG data from Experiment 1 was very different from previous findings on binocular combination using steady-state EEG with grating stimuli (Baker and Wade, 2017). One possible explanation is that the lower temporal frequency used here (2Hz, vs 5 or 7Hz in previous work) might be responsible for this difference. We therefore ran a second experiment to compare monocular and binocular responses at a range of temporal frequencies. Only EEG data were collected for this experiment, as the pupil response is negligible above around 2Hz (Spitschan et al., 2014); note that we originally chose 2Hz because it produces measurable signals for both EEG and pupillometry, yet is unfortunately optimal for neither.

Results from the temporal frequency experiment are shown in Figure 4. Figure 4a shows the Fourier spectra for responses to binocular flicker at 5 different frequencies (2, 4, 8, 16, and 30 Hz). From 2 to 16 Hz, clear signals are observed at each fundamental frequency, and typically also their higher harmonics (integer multiples of the fundamental). However, at 30 Hz (upper row), the responses recorded were not demonstrably above the noise baseline. Figure 4b compares the monocular and binocular responses at each stimulation frequency. Here we replicate the substantial summation effect across frequencies up to and including 16Hz (Fig. 4c), demonstrating that strong binocular facilitation in the EEG data of Experiment 1 cannot be attributed to our use of 2Hz flicker.

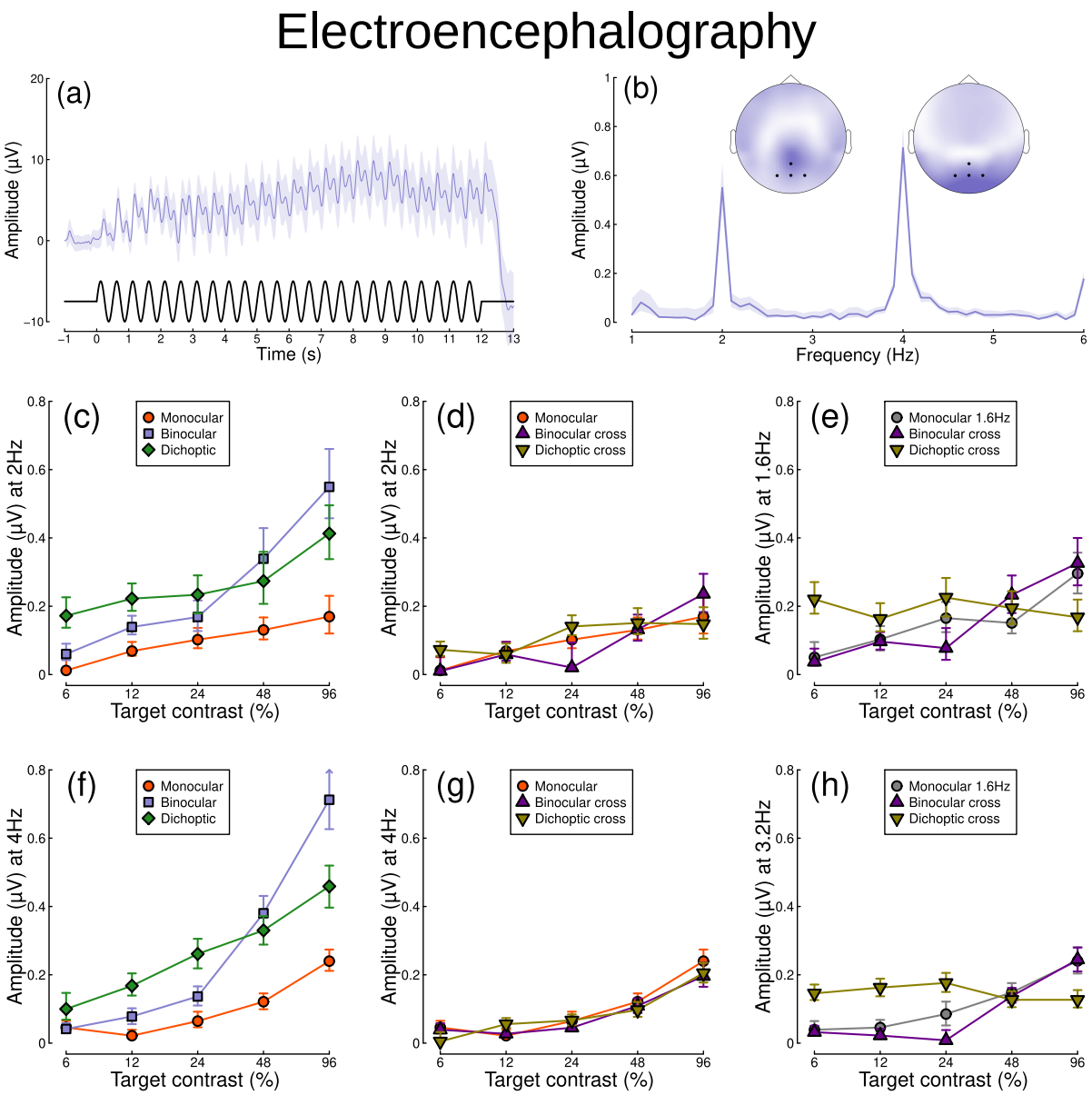

Figure 2: Summary of EEG results for N=30 participants. Panel (a) shows a group average waveform for binocular presentation (low pass filtered at 5Hz), with the driving signal plotted at the foot. Panel (b) shows the average Fourier spectrum, and inset scalp distributions. Black dots on the scalp plots indicate electrodes Oz, POz, O1 and O2. Panels (c,d) show contrast response functions at 2Hz for different conditions. Panel (e) shows contrast response functions at 1.6Hz for three conditions. Panels (f-h) are in the same format but for the second harmonic responses. Shaded regions and error bars indicate bootstrapped standard errors.

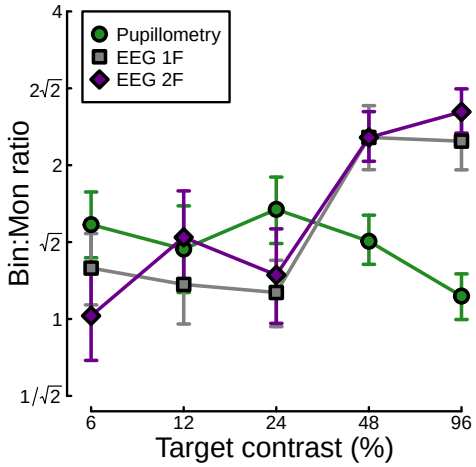

Figure 3: Ratio of binocular to monocular response for three data types. Each ratio is the average of ratios for N=30 participants, and error bars indicate bootstrapped standard errors.

4.3 Experiment 3

In Experiment 1 we found evidence of stronger binocular facilitation for cortical responses to luminance flicker (measured using EEG), compared with subcortical responses (measured using pupillometry). Since perception is dependent on cortical responses, these results provide a clear prediction for perceived contrast judgements indexed by psychophysical contrast matching paradigms (e.g. Anstis and Ho, 1998; Levelt, 1965; Quaia et al., 2018). We therefore conducted such an experiment, in which participants judged which of two stimuli had the greater perceived amplitude of flicker. One stimulus was a matching stimulus, that had a fixed binocular flicker amplitude of either 24% or 48% (temporal) contrast. The other stimulus was a target stimulus, the contrast of which was controlled by a staircase algorithm. We tested 9 ratios of contrast between the left and right eyes.

The results from the matching experiment are shown in Figure 5. Each data point indicates the contrast levels required in each eye that were perceptually equivalent to the binocular 24% (red circles) and 48% (blue circles) matching contrasts. At both matching contrasts, we see a very substantial increase in the physical contrast required for a monocular target (data points along the x- and y-axes), compared to a binocular target (points along the diagonal of x=y). For example with a 48% match, the monocular targets required contrasts close to 100%, whereas binocular targets required a contrast of around 50%. The data points between these extremes also fall close to the predictions of a linear summation model (diagonal dotted lines), and are inconsistent with a winner-takes-all (or MAX) model (dashed lines). Overall, these matching results are consistent with the approximately linear summation effects observed in the EEG data of Experiment 1 (Figure 2c,f).

4.4 Computational modelling

We fitted a computational model to the data from Experiments 1 & 3 using a hierarchical Bayesian approach. The model behaviour is displayed in Figure 6e-h, with empirical data replotted in Figure 6a-d for comparison. In general, the model captures the key characteristics of the empirical data. Note that there are some minor discrepancies, which are a consequence of the hierarchical nature of the modelling. In brief, the model is fitted

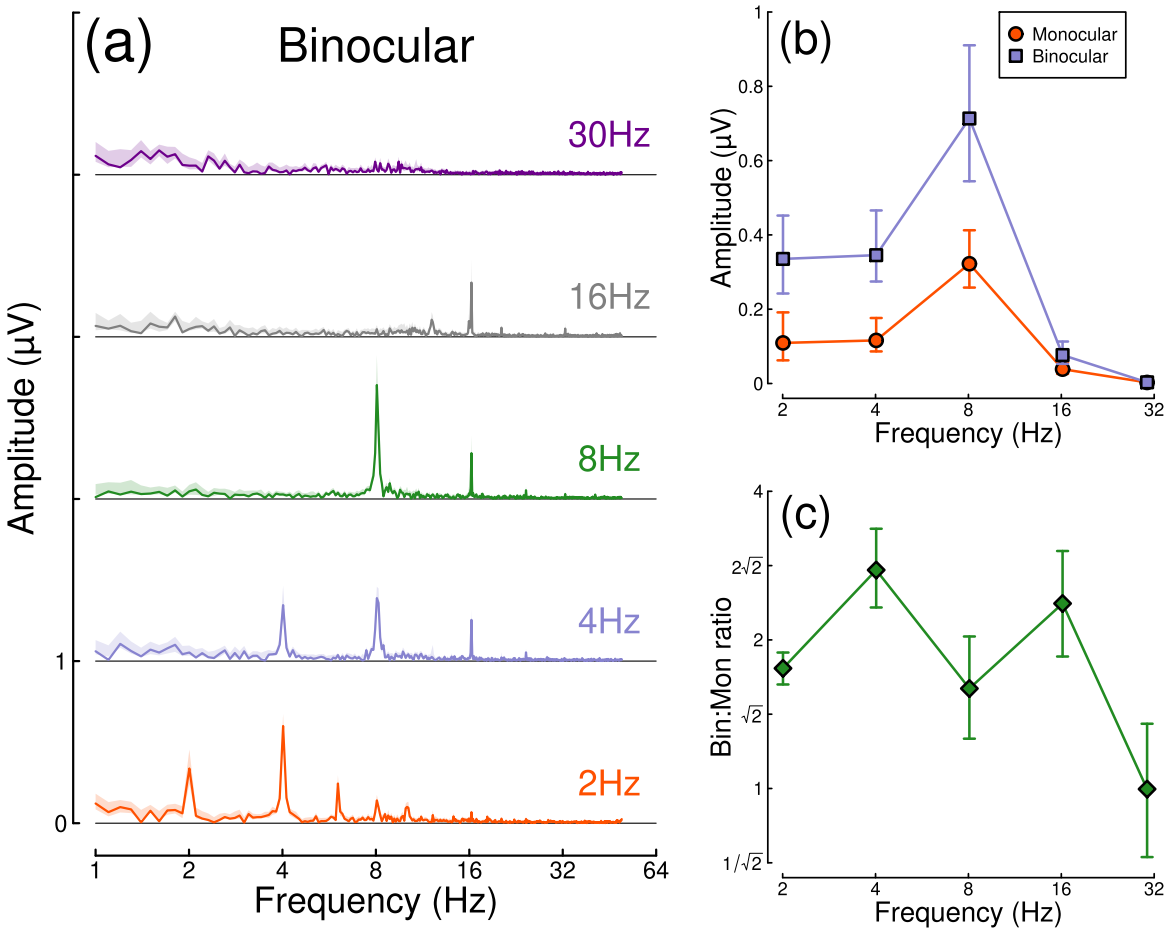

Figure 4: Binocular facilitation at different temporal frequencies. Panel (a) shows Fourier spectra for responses to binocular flicker at 5 different frequencies (offset vertically for clarity). Panel (b) shows the response at each stimulation frequency for monocular (red circles) and binocular (blue squares) presentation. Panel (c) shows the ratio of binocular to monocular responses. Error bars and shaded regions indicate bootstrapped standard errors across N=12 participants.

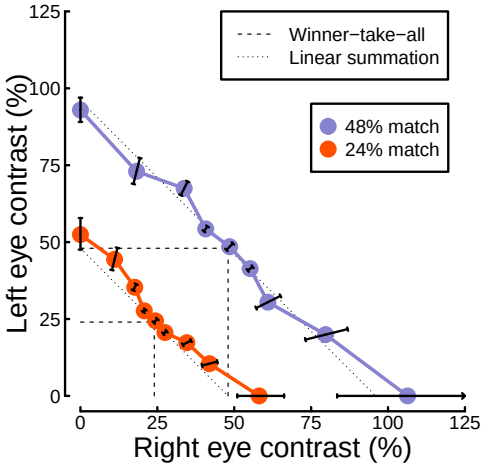

Figure 5: Contrast matching functions. Dotted and dashed lines are predictions of canonical summation models with a linear exponent (dotted) or an infinite exponent (dashed). Error bars indicate the standard error across participants (N=10), and are constrained along radial lines converging at the origin. Note that, for the 48% match, the data point on the x axis falls higher than 100% contrast. This is because the psychometric function fits for some individuals were interpolated such that the PSE fell above 100%, shifting the mean slightly above that value.

Table 1: Summary of median parameter values.

Data set	Z	k	w	Rmax
Pupillometry	2.37	0.01	0.67	0.00023
EEG 1F	2.59	0.15	0.02	0.0026
EEG 2F	2.21	0.06	0.01	0.00404
Matching	0.30	5.07	0.09	-

to the amplitudes for each participant, and group-level parameter estimates are derived based on these fits (see Table 1). This procedure discards the phase information, whereas the empirical averages are coherently averaged across participants (retaining phase information). This explains the amplitude differences between model and data, particularly at low target contrast levels, but is of little consequence for the pattern of relative responses across conditions, which is our main focus here.

We were particularly interested in comparing the weight of interocular suppression across data sets. We therefore plot the posterior distributions for this parameter for all four data sets (see Figure 6i). The key finding is that the pupillometry results (green distribution) display a much greater weight of interocular suppression compared with the other data sets (grey, purple and yellow distributions). There is no overlap between the pupillometry distribution and any of the other three. All four distributions are also meaningfully below a weight of 1 – the value that previous work using grating stimuli would predict (Baker and Wade, 2017; Meese et al., 2006), and the peak location of our prior distribution (black curve). These results offer an explanation of the empirical data: the strong interocular suppression for the pupillometry data is consistent with the weak binocular facilitation, and measurable dichoptic masking observed using that method. The weaker suppression for the other experiments is consistent with the near-linear binocular facilitation effects, and absent dichoptic masking.

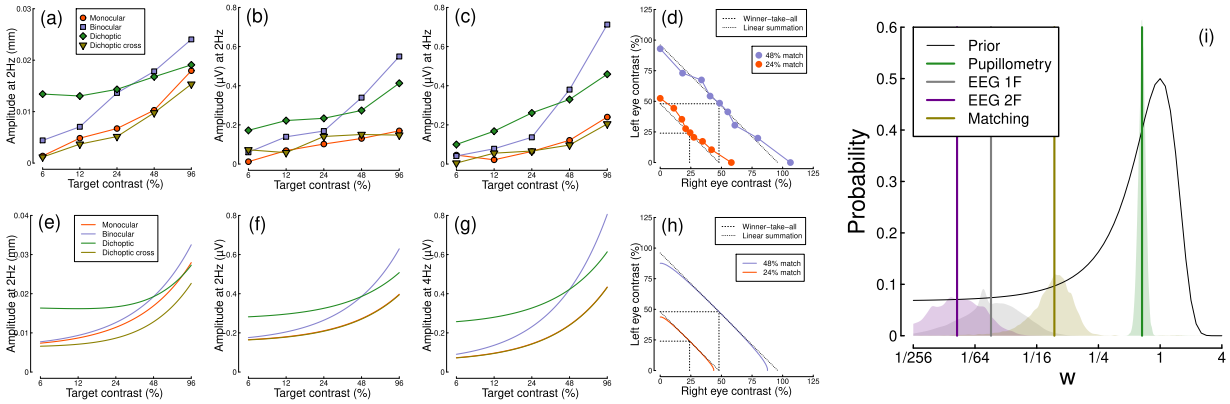

Figure 6: Summary of computational modelling. Panels (a-d) show empirical data from key conditions, replotted from earlier figures for the pupillometry (a), first harmonic EEG responses (b), second harmonic EEG responses (c) and contrast matching (d) experiments. Panels (e-h) show model behaviour for the same conditions, generated using the median group-level parameter values. Panel (i) shows the posterior probability distributions of the interocular suppression parameter for each of the four model fits. The pupillometry distribution (green) is centred about a substantially higher suppressive weight than for the other data types (note the logarithmic x-axis). The black curve shows the (scaled) prior distribution for the weight parameter.

5 Discussion

Using a novel paradigm that combines EEG and pupillometry, we found surprising results for the binocular integration of flickering light. In the visual cortex, the processing of spatially-uniform temporal luminance modulations seems to happen approximately linearly, with no evidence of interocular suppression. In the subcortical pathway, signal combination is more non-linear, with evidence of interocular suppression. This pattern of results was confirmed by computational modelling, which showed a much greater suppressive weight for the pupillometry data compared to the EEG data. Additionally, we found that perception of flickering light is consistent with a near-linear summation process, consistent with the cortical (EEG) responses.

The results of our main experiment were unexpected for both the pupillometry and the EEG measures. Previous studies investigating binocular combination of spatial patterns (i.e. sine wave grating stimuli) are all consistent with strong interocular suppression and weak binocular facilitation at high contrasts (Baker and Wade, 2017; Meese et al., 2006; Moradi and Heeger, 2009). Our second experiment ruled out the possibility that these differences were due to the lower temporal frequency (2Hz) used here. However, there is evidence of more extensive binocular facilitation for a range of other stimuli. For example, Quaia et al. (2018) observed a strong binocular facilitation (or 'supersummation') in the reflexive eye movement response to rapidly moving stimuli, and Spitschan and Cajochen (2019) report a similar result in archival data on melatonin suppression due to light exposure. In the auditory system, interaural suppression of amplitude modulation also appears to be weak when measured using a similar steady-state paradigm (Baker et al., 2020). Finally, psychophysical matching experiments using static stimuli also show near-linear behaviour for luminance increments (Anstis and Ho, 1998; Baker et al., 2012; Levelt, 1965), though not for luminance decrements (Anstis and Ho, 1998). Overall, this suggests that strong interocular normalization may be specific to spatial pattern vision, and not a general feature of sensory signal combination.

Given the above, where does this leave our understanding of the overarching purpose of signal combination? Baker and Wade (2017) point out that strong suppression between channels that are subsequently summed is equivalent to a Kalman filter, which is the optimal method for combining two noisy inputs (see also Ernst and Banks, 2002). This account has intuitive appeal, and is consistent with other models that propose binocular combination as a means of redundancy reduction (Li and Atick, 1994; May and Zhaoping, 2022). One

possibility is that optimal combination is useful for visual perception — a critical system for interacting with the local environment — and is therefore worth devoting the additional resource of inhibitory wiring between ocular channels. However the other examples given in the previous paragraph are primarily physiological responses (pupil size, eye movements, hormone release) that may benefit less from an increased signal-to-noise ratio, or otherwise be phylogenetically older than binocular pattern vision. Conceptualised another way, the brain can repurpose a generic architecture for different situational demands by adjusting parameter values (here the weight of interocular suppression) to achieve different outcomes. Our future work in this area intends to compare binocular combination for specific photoreceptor pathways, including different cone classes, and intrinsically photoreceptive retinal ganglion cells.

Pupil size determines the total amount of light falling on the retina. It is therefore the case that fluctuations in pupil diameter have a downstream effect on the signals reaching cortex. We did not incorporate such interactions into our computational model, though in principle this might be worthwhile. However we anticipate than any such effects would be small, since pupil modulations at 2Hz are in the order of 2% of overall diameter (e.g. Spitschan et al., 2014). It is also the case that cortical activity can modulate pupil diameter, usually through arousal and attention mechanisms (e.g. Bradley et al., 2008). We think it unlikely that these temporally coarse processes would have a differential effect on e.g. monocular and binocular stimulation conditions in our experiment, and any fluctuations during an experimental session (perhaps owing to fatigue) will be equivalent for our comparisons of interest. Therefore we make the simplifying assumption that the two pathways are effectively distinct, but hope to investigate this directly in future neuroimaging work.

Classic studies investigating the neurophysiological architecture of V1 reported that cells in cytochrome-oxidase 'blobs' (Horton and Hubel, 1981; Livingstone and Hubel, 1984) are biased towards low spatial frequencies (Edwards et al., 1995; Tootell et al., 1988), and relatively insensitive to stimulus orientation (Horton and Hubel, 1981; Livingstone and Hubel, 1984; though see Economides et al., 2011). As the blob regions are embedded within ocular dominance columns (Horton and Hubel, 1981), they are also largely monocular (Livingstone and Hubel, 1984; Tychsen et al., 2004). More recent work has reported psychophysical evidence for unoriented chromatic (Gheiratmand et al., 2013) and achromatic (Meese and Baker, 2011) mechanisms, that also appear to be monocular. Our use of luminance flicker might preferentially stimulate these mechanisms, perhaps explaining why our EEG data show little evidence of binocular interactions. Indeed, our EEG results could potentially be explained by a model involving entirely non-interacting monocular channels, with the binocular facilitation effects we find (e.g. Figures 3 & 4) owing to additivity of the electrophysiological response across independent monocular cells, rather than requiring binocular neurons. However our matching data (Figure 5), as well as everyday perceptual experience, indicates that luminance signals must be combined physiologically at some stage, and that this process still involves weak or absent interocular suppression.

6 Conclusions

We have demonstrated that binocular combination of flickering light differs between cortical and subcortical pathways. Flicker was also associated with substantially weaker interocular suppression, and stronger binocular facilitation, compared to combination of spatial luminance modulations in visual cortex. Our computational framework for understanding signal combination permits direct comparisons between disparate experimental paradigms and data types. We anticipate that this will help elucidate the constraints the brain faces when combining different types of signals to govern perception, action and biological function.

7 Acknowledgements

Supported by Biotechnology and Biological Sciences Research Council grant BB/V007580/1 awarded to DHB and ARW, and Wellcome Trust grant 213616/Z/18/Z to AB.

8 Author contributions

Federico Segala: Methodology, software, formal analysis, investigation, data curation, writing - original draft, writing - review & editing, visualization. **Aurelio Bruno**: Conceptualization, writing - review & editing, supervision, project administration, funding acquisition. **Myat Aung**: Software, resources, writing - review & editing. **Alex Wade**: Conceptualization, methodology, resources, writing - review & editing, supervision, project administration, funding acquisition. **Daniel Baker**: Conceptualization, methodology, software, formal analysis, investigation, resources, data curation, writing - original draft, writing - review & editing, visualization, supervision, project administration, funding acquisition.

9 Declaration of interests

The authors declare no competing interests.

References

Angée C, Nedelec B, Erjavec E, Rozet J-M, Fares Taie L. 2021. Congenital microcoria: Clinical features and molecular genetics. *Genes (Basel)* **12**. doi:10.3390/genes12050624

Anstis S, Ho A. 1998. Nonlinear combination of luminance excursions during flicker, simultaneous contrast, afterimages and binocular fusion. *Vision Res* **38**:523–39. doi:10.1016/s0042-6989(97)00167-3

Baker DH. 2021. Statistical analysis of periodic data in neuroscience. *Neurons, Behavior, Data analysis, and Theory* **5**. doi:10.51628/001c.27680

Baker DH, Lygo FA, Meese TS, Georgeson MA. 2018. Binocular summation revisited: Beyond $\sqrt{2}$. *Psychol Bull* **144**:1186–1199. doi:10.1037/bul0000163

Baker DH, Meese TS, Georgeson MA. 2007. Binocular interaction: Contrast matching and contrast discrimination are predicted by the same model. *Spat Vis* **20**:397–413. doi:10.1163/156856807781503622

Baker DH, Vilidaite G, McClarnon E, Valkova E, Bruno A, Millman RE. 2020. Binaural summation of amplitude modulation involves weak interaural suppression. *Sci Rep* **10**:3560. doi:10.1038/s41598-020-60602-5

Baker DH, Wade AR. 2017. Evidence for an optimal algorithm underlying signal combination in human visual cortex. *Cereb Cortex* **27**:254–264. doi:10.1093/cercor/bhw395

Baker DH, Wallis SA, Georgeson MA, Meese TS. 2012. Nonlinearities in the binocular combination of luminance and contrast. *Vision Res* **56**:1–9. doi:10.1016/j.visres.2012.01.008

Barrionuevo PA, Cao D. 2016. Luminance and chromatic signals interact differently with melanopsin activation to control the pupil light response. *J Vis* **16(11)**:29. doi:10.1167/16.11.29

Barrionuevo PA, Nicandro N, McAnany JJ, Zele AJ, Gamlin P, Cao D. 2014. Assessing rod, cone, and melanopsin contributions to human pupil flicker responses. *Invest Ophthalmol Vis Sci* **55**:719–27. doi:10.1167/iovs.13-13252

Bradley MM, Miccoli L, Escrig MA, Lang PJ. 2008. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* **45**:602–7. doi:10.1111/j.1469-8986.2008.00654.x

Campbell FW, Green DG. 1965. Monocular versus binocular visual acuity. *Nature* **208**:191–2. doi:10.1038/208191a0

Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A. 2017. Stan: A probabilistic programming language. *Journal of Statistical Software* **76**:1–32. doi:10.18637/jss.v076.i01

Delorme A, Makeig S. 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* **134**:9–21. doi:10.1016/j.jneumeth.2003.10.009

Ding J, Sperling G. 2006. A gain-control theory of binocular combination. *Proc Natl Acad Sci U S A* **103**:1141–6. doi:10.1073/pnas.0509629103

Doesschate J ten, Alpern M. 1967. Effect of photoexcitation of the two retinas on pupil size. *J Neurophysiol* **30**:562–76. doi:10.1152/jn.1967.30.3.562

Economides JR, Sincich LC, Adams DL, Horton JC. 2011. Orientation tuning of cytochrome oxidase patches in macaque primary visual cortex. *Nat Neurosci* **14**:1574–80. doi:10.1038/nn.2958

Edwards DP, Purpura KP, Kaplan E. 1995. Contrast sensitivity and spatial frequency response of primate cortical neurons in and around the cytochrome oxidase blobs. *Vision Res* **35**:1501–23. doi:10.1016/0042-6989(94)00253-i

Ernst MO, Banks MS. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**:429–33. doi:10.1038/415429a

Figueira JSB, Kutlu E, Scott LS, Keil A. 2022. The freqtag toolbox: A principled approach to analyzing electrophysiological time series in frequency tagging paradigms. *Dev Cogn Neurosci* **54**:101066. doi:10.1016/j.dcn.2022.101066

Gheiratmand M, Meese TS, Mullen KT. 2013. Blobs versus bars: Psychophysical evidence supports two types of orientation response in human color vision. *Journal of Vision* **13(1)**:2. doi:10.1167/13.1.2

Horton JC, Hubel DH. 1981. Regular patchy distribution of cytochrome oxidase staining in primary visual cortex of macaque monkey. *Nature* **292**:762–4. doi:10.1038/292762a0

Hubel DH, Wiesel TN. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* **160**:106–54. doi:10.1113/jphysiol.1962.sp006837

Kassner M, Patera W, Bulling A. 2014. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction.Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. UbiComp '14 Adjunct; ACM. pp. 1151–1160. doi:10.1145/2638728.2641695

Kingdom FAA, Libenson L. 2015. Dichoptic color saturation mixture: Binocular luminance contrast promotes perceptual averaging. *J Vis* **15(5)**:2. doi:10.1167/15.5.2

Legge GE. 1984. Binocular contrast summation–ii. Quadratic summation. *Vision Res* **24**:385–94. doi:10.1016/0042-6989(84)90064-6

Levelt WJ. 1965. Binocular brightness averaging and contour information. *Br J Psychol* **56**:1–13. doi:10.1111/j.2044-8295.1965.tb00939.x

Li Z, Atick JJ. 1994. Efficient stereo coding in the multiscale representation. *Network: Computation in Neural Systems* **5**:157–174. doi:10.1088/0954-898X_5_2_003

Livingstone MS, Hubel DH. 1984. Anatomy and physiology of a color system in the primate visual cortex. *J Neurosci* **4**:309–56. doi:10.1523/JNEUROSCI.04-01-00309.1984

Mathôt S. 2018. Pupillometry: Psychology, physiology, and function. *J Cogn* **1**:16. doi:10.5334/joc.18

May K, Zhaoping L. 2022. Li and Atick's theory of efficient binocular coding: A tutorial and mini-review. *Vision Res* **201**:107950. doi:10.1016/j.visres.2021.08.005

McDougal DH, Gamlin PD. 2008. Pupillary control pathways.The Senses: A Comprehensive Reference. Elsevier. pp. 521–536.

Meese TS, Baker DH. 2011. A reevaluation of achromatic spatio-temporal vision: Nonoriented filters are monocular, they adapt, and can be used for decision making at high flicker speeds. *iPerception* **2**:159–82. doi:10.1068/i0416

Meese TS, Georgeson MA, Baker DH. 2006. Binocular contrast vision at and above threshold. *J Vis* **6**:1224–43. doi:10.1167/6.11.7

Moradi F, Heeger DJ. 2009. Inter-ocular contrast normalization in human visual cortex. *J Vis* **9(3)**:13. doi:10.1167/9.3.13

Norcia AM, Appelbaum LG, Ales JM, Cottereau BR, Rossion B. 2015. The steady-state visual evoked potential in vision research: A review. *J Vis* **15(6)**:4. doi:10.1167/15.6.4

Purves D, Brannon EM, Cabeza R, LaBar KS, Huettel SA, Platt ML, Woldorff MG. 2008. Principles of Cognitive Neuroscience. Oxford University Press.

Quaia C, Optican LM, Cumming BG. 2018. Binocular summation for reflexive eye movements. *J Vis* **18(4)**:7. doi:10.1167/18.4.7

Schrödinger E. 1926. Die gesichtsempfindungen.Mueller-Pouillets Lehrbuch Der Physik (11th Ed.). Vieweg: Braunschweig. pp. 456–560.

Spitschan M, Cajochen C. 2019. Binocular facilitation in light-mediated melatonin suppression? *J Pineal Res* **67**:e12602. doi:10.1111/jpi.12602

Spitschan M, Jain S, Brainard DH, Aguirre GK. 2014. Opponent melanopsin and s-cone signals in the human pupillary light response. *Proc Natl Acad Sci U S A* **111**:15568–72. doi:10.1073/pnas.1400942111

Tootell RB, Silverman MS, Hamilton SL, Switkes E, De Valois RL. 1988. Functional anatomy of macaque striate cortex. V. Spatial frequency. *J Neurosci* **8**:1610–24. doi:10.1523/JNEUROSCI.08-05-01610.1988

Tychsen L, Wong AM-F, Burkhalter A. 2004. Paucity of horizontal connections for binocular vision in v1 of naturally strabismic macaques: Cytochrome oxidase compartment specificity. *The Journal of Comparative Neurology* **474**:261–275. doi:10.1002/cne.20113

Victor JD, Mast J. 1991. A new statistic for steady-state evoked potentials. *Electroencephalogr Clin Neurophysiol* **78**:378–88. doi:10.1016/0013-4694(91)90099-p

Wang C-A, Munoz DP. 2015. A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Curr Opin Neurobiol* **33**:134–40. doi:10.1016/j.conb.2015.03.018

Wyatt HJ, Musselman JF. 1981. Pupillary light reflex in humans: Evidence for an unbalanced pathway from nasal retina, and for signal cancellation in brainstem. *Vision Res* **21**:513–25. doi:10.1016/0042-6989(81)90097-3